An Investigation of the role of the *S. cerevisiae* RNA-binding protein Scp160p/vigilin in the aggregation of Q/N-rich proteins

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Matthew Hong Kei Cheng

aus Hongkong

Tübingen

2018

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS AND THANKS

ZUSAMMENFASSUNG

RNA-Bindeproteine der Vigilin Familie stellen eine von Hefe bis zu Säugetieren hochkonservierte Proteinfamilie dar. Funktionell konnte in verschiedenen Organismen eine Verbindung zwischen Vigilin und einer Vielzahl zellulärer Prozesse gezeigt werden. Dies beinhaltet sowohl cytoplasmatische als auch nukleäre Prozesse, zu denen unter anderem Ploidiewartung und Bildung von Heterochromatin sowie die Regulation der Entstehung von processing bodies, mRNA-Transport, -Translation, und -Stabilität zählen. So konnten frühere Arbeiten unserer Forschungsgruppe zeigen, dass das Vigilin Homolog von *Saccharomyces cerevisiae*, Scp160p, die Wiederverwendung von tRNAs abhängig von Codonzusammensetzung begünstigt.

Diese Arbeit erforscht diese Rolle weiter, indem sie untersucht, wie der Verlust von Scp160p die Biosynthese von aggregationsanfälligen Polyglutamin (PolyQ)-Reporterproteinen, welche sich in ihrer Codonverwendugn unterscheiden, beeinflusst. Mikroskopie und biochemische Analysen zeigen, dass die Aggregation der PolyQ-Reporterproteine in Zellen mit deletiertem Scp160p (*scp160Δ* Zellen) unabhängig von der Codonverwendung reduziert ist. Weitergehend wurde dies durch eine Kombination an Filter-trap Bindung, welches SDS-resistente aggregierte Proteine einfängt, mit quantitativer Massenspektrometrie durch Dimethyl-Labeling, für das endogene *S. cerevisiae* Proteom untersucht. Die Analyse der Daten zeigte, dass die Aggregation vieler polyQ- und Glutamin/Asparagin-reicher (Q/N-reicher) Proteine, darunter der Transkriptionsfaktor Cyc8p und das RNA-Bindeprotein Nab3p, in Abwesenheit von Scp160p reduziert ist.

Vorläufige Ergebnisse deuten darauf hin, dass der Verlust von Scp160p die Translationskinetik stören könnte, wodurch co-translationale Faltung und die Wahrscheinlichkeit der Aggregation von Proteinen beeinflusst wird. Durch die Modulation der Q/N-vermittelten Proteinaggregation könnte Scp160p somit viele zelluläre Prozesse beeinflussen.

SUMMARY

The vigilin family of RNA-binding proteins are highly conserved from yeast to mammals. Work from various organisms have linked vigilin to a multitude of cellular processes both in the nucleus and cytoplasm. These processes include ploidy maintenance, heterochromatin formation, as well as regulation of processing-body formation, mRNA localization, translation, and stability. Our lab has previously reported a role for the *Saccharomyces cerevisiae* vigilin homolog, Scp160p, in promoting the reuse of tRNAs in the context of codon composition.

This work explores this role further by assessing how loss of Scp160p affects biosynthesis of aggregation-prone polyglutamine (polyQ) reporters that differ in their codon usage. Microscopy and biochemical analysis show that aggregation of the polyQ reporters is reduced in *scp160Δ* cells, regardless of codon usage. This effect on protein aggregation was further assayed for the endogenous *S. cerevisiae* proteome by combining filter-trap binding, to capture SDS-resistant aggregated proteins, and quantitative mass spectrometry by dimethyl labelling. Such an analysis revealed that aggregation of many polyQ and glutamine/asparagine-rich (Q/N-rich) proteins, including the transcription factor Cyc8p and RNA-binding protein Nab3p, were reduced in the absence of Scp160p. Additional work preliminarily suggests that loss of Scp160p may perturb translation kinetics, thereby influencing co-translational folding and likelihood of protein aggregation. In this way, Scp160p may influence many cellular processes by modulating Q/N-mediated protein aggregation.

ABBREVIATIONS AND SYMBOLS

4EHP – eIF4E homologous protein

aa-tRNA – aminoacylated tRNA

AIN-2 –ALG-1 interacting protein

ASC1/Asc1p –Absence of growth suppressor of Cyp1 gene/protein

ASH1/Ash1p – Asymmetric synthesis of HO gene/protein

BFR1/Bfr1p –Brefeldin A resistance gene/protein

BNI1– Bud neck involved gene

BUD6/Bud6p – BUD site selection gene/protein

C. elegans – Caenorhabditis elegans

C9ORF72 – Chromosome 9 open reading frame 72

CAF1/Caf1p – CCR4-associated factor gene/protein

CCR4-NOT – Carbon catabolite repression/Negative on TATA

CDC19/Cdc19p, CDC48/Cdc48p – Cell division cycle 19/48 gene/protein

ChIP – Chromatin immunoprecipitation

CHX – Cycloheximide

CLEM – Correlative light and electron microscopy

CLN3 – Cyclin gene

C_t – Cycle threshold

CTCF – CCCTC-binding factor

CYC8/Cyc8p – Cytochrome C gene/protein

DBD – DNA-binding domain

DCP1, *DCP2*/Dcp1p, Dcp2p – mRNA decapping gene/protein

DDP1 – Dodecasatellite-binding protein

DEPC – Diethyl pyrocarbonate

DHH1/Dhh1p – DEAD-box helicase homolog gene/protein

DNA – Deoxyribonucleic acid

DOM34/Dom34p –Duplication of multilocus region gene/protein

dPER – *Drosophila* PERIOD

Drosophila S2 cells – *Drosophila* Schneider 2 cells

dsRBD – double-stranded RNA-binding domain

E. coli – Escherichia coli

EAP1/Eap1p –eIF4E-associated protein gene/protein

EDC3/Edc3p – Enhancer of mRNA decapping gene/protein

EDTA – Ethylenediaminetetraacetic acids

eEF1A – Eukaryotic elongation factor 1A

eIF – Eukaryotic initiation factor

ER – Endoplasmic reticulum

ERAD – ER-associated degradation

EWS – Ewings sarcoma

FACS – Fluorescence assisted cell sorting

FRQ – Frequency

FUS – Fused in sarcoma

GFP – Green fluorescent protein

HA – Hemagglutinin

HBS1/Hbs1p –Hsp70 subfamily B suppressor gene/protein

HD – Huntington's disease

HDLBP – High-density lipoprotein binding protein

hnRNPA1 – Heterogeneous nuclear ribonucleoprotein A1

HSP – Heat shock protein

HTT/Htt – Huntingtin gene/protein

Igf2 – Insulin-like growth factor 2

INQ – "Intranuclear quality control" compartment

IPOD – Insoluble protein deposit

IRE1/Ire1p – Inositol-requiring enzyme gene/protein

JUNQ – "Juxtanuclear quality control" compartment

KH – hnRNP-K Homology

LCR – Low-complexity region

LSM1-7/Lsm1-7p – Like Sm gene/protein

LTN1/Ltn1p – Listerin gene/protein

MDR1/Mdr1p – Multidrug resistance gene/protein

MHC – Major histocompatibility complex

m*HTT*/mHtt – CAG-expanded Huntingtin gene/protein

MICB –MHC Class I-related chain B

miRISC – microRNA effector RNA-induced silencing complex

miRNA – microRNA

mRNA – Messenger RNA

mRNP – Messenger ribonucleoprotein

MYO4/Myo4p – Myosin gene/protein

MZT – Maternal-to-zygotic transition

NAB3/Nab3p – Nuclear polyadenylated RNA-binding gene/protein

NEW1/New1p – Nu+ gene/protein

NGD – No-go decay

NK cells – Natural killer cells

NKG2D – Natural killer group 2 member D

NPL4/Npl4p – Nuclear protein localization gene/protein

NSD – Non-stop decay

P-body – Processing body

PABP – Poly(A)-binding protein

PAN2, *PAN3*/Pan2p, Pan3p – Poly(A)-binding protein-dependent poly(A)
    ribonuclease gene/protein

PARN – Poly(A) ribonuclease

PBS – Phosphate buffered saline

PBS-T – Phosphate buffered saline + 0.1% Tween

PGK1/Pgk1p –Phosphoglucose kinase gene/protein

PI – Propidium iodide

POL – Polarity and secretion factor

Poly(A) - Polyadenosine

PolyN – Polyasparagine

PolyQ – Polyglutamine

POM34/Pom34p –Pore membrane gene/protein

POP2/Pop2p – PGK promoter overproduction gene/protein

PrD – Prion-like domain

PRF – Programmed ribosomal frameshift

PRR – Proline-rich region

PTR – Post-transcriptional regulation

Q/N – Glutamine/Asparagine

RBD – RNA-binding domain

RBP – RNA-binding protein

REST – mRNA-encoded slowdown of translation

RNA – Ribonucleic acid

RPL/Rpl – Ribosomal protein of the large subunit gene/protein

RPS/Rps – Ribosomal protein of the small subunit gene/protein

RQC – Ribosome-associated quality control

RQC1/Rqc1p*, RQC2*/Rqc2p – Ribosome quality control complex gene/protein

RRM – RNA-recognition motif

RT-qPCR – Reverse transcription quantitative polymerase chain reaction

S. cerevisiae – *Saccharomyces cerevisiae*

S. pombe – *Schizosaccharomyces pombe*

SCP160/Scp160p –*S. cerevisiae* protein involved in the control of ploidy
gene/protein

SESA – A protein network consisting of Smy2p, Eap1p, Scp160p, and Asc1p

SG – Stress granule

SIS1/Sis1p – *SIT4* suppressor gene/protein

SKI/Ski – Superkiller genes/protein complex

SMY2/Smy2p –Suppressor of *myo2-66* gene/protein

SNP – Single nucleotide polymorphism

SPA2 – Spindle pole antigen gene

SPB – Spindle pole body

SRP – Signal recognition particle

SS – Signal sequence

SSA1/Ssa1p – Stress-seventy subfamily A gene/protein

STI1/Sti1p – Stress inducible gene/protein

SUP35/Sup35p – Suppressor gene/protein

TAF15 – TATA-binding protein associated factor 15

TIA-1 – T-cell-restricted intracellular antigen 1

TM – Transmembrane domain

tRNA – Transfer RNA

UFD1/Ufd1p – Ubiquitin fusion degradation gene/protein

UPR – Unfolded protein response

Usp2 – Ubiquitin specific peptidase

UTR – Untranslated region

XRN1/Xrn1p – Exoribonuclease gene/protein

Vgl1 – Vigilin

YDJ1/Ydj1p – Yeast DNAJ1 gene/protein

x

1. INTRODUCTION

1.1 – Regulating gene expression via post-transcriptional mechanisms

Post-transcriptional regulation (PTR) is an essential aspect of gene expression control occurring at the RNA level. It encompasses all processes of RNA metabolism including the nuclear steps of 5' cap addition, polyadenylation, splicing, and nuclear export. In the cytoplasm, mature mRNAs can be regulated on three primary levels which impact localization, translation, and/or stability. PTR presents several advantages over transcription-based gene expression control. First, PTR occurs rapidly and allows cells respond to stimuli in a critical and timely manner. Second, in the case of mRNA localization and translation, this regulation is reversible and allows efficient restoration of cell functions after stimuli dissipation. Thirdly, the spatial restriction of mRNAs facilitates a high rate of local protein production and accumulation which prevents potentially detrimental off-site protein function. This is especially important for large cells like neurons and for transcriptionally silent cell types like enucleated erythrocytes and early embryos.

PTR is primarily mediated by RNA-binding proteins (RBPs), which bind target transcripts via a specific sequence and/or structural element. These RNA elements are found largely in mRNA 3'-untranslated regions (UTRs), but also in 5'-UTRs and open reading frames. RNA-binding is mediated by RNA-binding domains (RBDs) of RBPs, most of which fall into 4 canonical classes: RNA recognition motif (RRM), double-stranded RNA-binding domain (dsRBD), hnRNP K-homology (KH) domain, and zinc fingers (Cléry and Allain, 2012). However, many proteins have now been found to bind RNA without one of these canonical RBDs (Hentze et al., 2018). Nevertheless, RNA-binding may be modulated by N- and C-terminal RBD extensions and/or cooperativity. As such, RBDs are often present in tandem repeats in RBPs (Cléry and Allain, 2012; Lunde et al., 2007; Nicastro et al., 2015). Moreover, multiple RBPs may bind a single mRNA to form distinct messenger ribonucleoprotein (mRNP) complexes and confer combinatorial regulation (Hogan et al., 2008).

1.1.1 – Translation control coordinates production

Translation initiation of an mRNA involves the assembly of the initiation complex (eIF4F) at the mRNA 5' cap, which includes of the cap-binding eIF4E and the scaffold eIF4G (Sonenberg and Hinnebusch, 2009). Since initiation is often rate-

limiting, it is therefore not surprising that this step is a target for regulation (Pichon et al., 2012). One mechanism to block translation initiation is to prevent the assembly of eIF4F by binding eIF4E to preclude recruitment of additional eIFs. The *Drosophila* eIF4E-binding protein Cup has been described to function in this capacity in early oocytes and embryos (Lasko, 2012; Nelson et al., 2007). Alternatively, eIF4E itself can be prevented from binding the 5' cap by competition from the eIF4E homologous protein 4EHP (Jackson et al., 2010).

Translation of *ASH1* mRNA in *Saccharomyces cerevisiae (S. cerevisiae)* is similarly repressed by the RBP Khd1p, which interacts with eIF4G1 (Paquin et al., 2007). Another RBP Puf6p, acts parallel to Khd1p and blocks another stage of initiation by preventing 60S ribosomal subunit joining (Deng et al., 2008). Translation repression of *ASH1* mRNA at the distal bud tip is relieved by phosphorylation of Khd1p and Puf6p, leading to their dissociation from the mRNA (Deng et al., 2008; Paquin et al., 2007).

The step of initiation is also targeted in response to stresses such as amino acid starvation, through phosphorylation of eIF2α and inactivation of 43S scanning (Sonenberg and Hinnebusch, 2009). In this way, genes whose translation are cap-independent continue to be expressed during stress despite a global repression of cap-dependent translation.

Translation regulation can also function through the poly(A) tail, which can facilitate circularization of mRNAs to promote translation (Sonenberg and Hinnebusch, 2009). This occurs via poly(A) binding proteins (PABP) on the poly(A) tail which mediate a 5' cap-eIF4E-eIF4G-PABP-poly(A) tail interaction. Circularization is dependent on the length of the poly(A) tail and so can be regulated by poly(A) shortening or lengthening by deadenylases and poly(A) polymerases, respectively (Hinnebusch and Lorsch, 2012; Tadros and Lipshitz, 2005).

1.1.2 – Regulating mRNA stability to clear unwanted messages

A more permanent form of PTR acts at the level of mRNA stability, which targets the 5' cap and the 3' poly(A) tail. Destabilization of an mRNA begins with shortening of the poly(A) tail by deadenylases. In *S. cerevisiae* and *Drosophila*, they include the major and minor cytoplasmic deadenylation complexes, CCR4-NOT and PAN2-PAN3, respectively (Parker, 2012). A third deadenylase, PARN, is found in many animals including *Xenopus* (Semotok et al., 2005). It is thought that PAN2-PAN3 mediates

2

initial deadenylation of the poly(A) tail to a length of ~65 nucleotides, after which CCR4-NOT takes over (Parker, 2012). Once the poly(A) tail is removed, the mRNA can be degraded in the 3' to 5' direction by the exosome.

Alternatively, a sufficiently shortened poly(A) tail (~10-12 nucleotides in *S. cerevisiae*) can become a substrate for binding of the Lsm1-7p complex. This complex works in concert with other factors to inhibit translation and stimulate decapping by the Dcp1p/Dcp2p complex (Parker, 2012). Loss of the 5' cap exposes the mRNA to 5' to 3' decay by Xrn1p. Notably during stress, mRNAs associated with decapping factors and Xrn1p are localized to processing-bodies (P-bodies), sites of mRNA decay and/or storage (Decker and Parker, 2012). In accordance, P-body formation is dependent on a self-interacting Yjef-N domain and Q/N-rich region in the decapping activators Edc3p and Lsm4p, respectively (Decker et al., 2007).

1.1.3 –mRNA transport puts things in the right place

mRNA localization depends on *cis*-element(s) in the mRNA called zip-codes or localization elements, which determine bound RBPs and thereby destination (Jansen and Niessing, 2012). This can lead to active transport, where mRNPs are directed along microtubule or actin filaments by motor proteins. In *Drosophila* oogenesis, the major morphogens in body patterning – *oskar*, *bicoid*, and *gurken* mRNAs – are transported into the oocyte by the dynein motor protein toward the microtubule minus end (Lasko, 2012). Interestingly, transport of *oskar* mRNA inside the oocytes switches to the kinesin motor protein toward the microtubule plus end.

In contrast, mRNA localization in *S. cerevisiae* occurs along actin filaments and utilizes the myosin motor protein Myo4p (Singer-Krüger and Jansen, 2014). This has been well characterized for the localization of *ASH1* mRNA to the distal bud tip, where its expression suppresses mating type switching in the daughter cell (Jansen et al., 1996). Moreover, a subset of mRNAs encoding polarity and secretion factors (POLs) are co-transported with the cortical endoplasmic reticulum (ER) to the incipient bud site or mating projection (Fundakowski et al., 2012; Gelin-Licht et al., 2012). Here, association with the membrane is mediated by the RBP of the mRNP and transport by Myo4p.

At their destination, mRNAs are kept in place by anchoring. *oskar* mRNA at the *Drosophila* posterior pole plasm is held in place by the long isoform of Oskar protein, which stimulates endocytosis and rearrangement of the F-actin cytoskeleton (Tanaka

et al., 2011; Vanzo and Ephrussi, 2002). Conversely, anchoring of *ASH1* mRNA in *S. cerevisiae* to the distal bud tip requires the actin-interacting proteins Bni1p and Bud6p, remodeling of the mRNP, and translation activation of the mRNA (Singer-Krüger and Jansen, 2014). Anchoring is also important for mRNAs whose localization depends on a passive diffusion and capture mechanism (Lasko, 2012).

1.1.4 – Additional mechanisms to regulate mRNAs

The translation and stability of mRNAs can be further modulated by non-coding RNAs such as microRNAs (miRNA) and long non-coding RNAs (lncRNA) (Bartel, 2018; Yoon et al., 2013). In the case of miRNA-mediated regulation, the level of base-complementarity of the miRNA seed region with the mRNA target site determines whether the mRNA is translational repressed or degraded (Bartel, 2018).

Another interesting layer to PTR occurs with modifications of the RNA itself. One example of this is m^{6}A RNA methylation, which is a reversible modification conserved from bacteria to humans (Yue et al., 2015). m^{6}A methylation is enriched in non-coding regions and leads to mRNA translation inhibition and destabilization. Another example is A-to-I editing, in which adenosines are deaminated to inosines (read as guanosine). This is conserved only in animals and filamentous *ascomycetes*, and is carried out by adenosine-deaminases acting on RNAs (ADAR) proteins (Bian et al., 2018). Editing can not only alter the protein product sequence and function, but also influence mRNA translation and stability, as well as the miRNA pathway (Licht and Jantsch, 2016; Liu et al., 2014a).

1.2 – The impact of codon usage on gene expression control

Although the regulation of translation initiation is discussed above, elongation can also have a great impact on protein biosynthesis. A major determinant of translation elongation kinetics is the codon usage of mRNAs. As there are more combinations of nucleotide triplets than standard amino acids (61 different codons for 20 standard amino acids), a single amino acid may be encoded by more than one synonymous codon. Notably, there is a codon usage bias such that codons encoding the same amino acid (synonymous codons) are not used equally in coding regions. Furthermore, the usage frequency is of a codon is mirrored by the cellular abundance of its cognate tRNA (Ikemura, 1985). Finally, because the third nucleotide of a codon can engage in non-Watson/Crick base-pairing (to "wobble"), a tRNA species can

decode more than one codon (Crick, 1966; Ikemura, 1985). These parameters have coevolved and together give rise to codon optimality, a metric for the quality of codons in an mRNA (Ikemura, 1985).

Analyses across organisms have shown that codon usage and optimality vary not only between genomes but also within a genome (Bennetzen and Hall, 1982; Duret, 2000; Gouy and Gautier, 1982; Ikemura, 1985). Distinct codon usage profiles can help coordinate gene expression. In particular, highly expressed genes tend to be enriched for optimal codons which are translated faster and with higher fidelity (Bazzini et al., 2016; Ikemura, 1985). This is believed to be due to the higher availability of cognate and isoaccepting tRNAs (Fluitt et al., 2007; Ikemura, 1985; Spencer et al., 2012). Moreover, genes that encode components of the same pathway or complex also have similar codon usage, which can help match their relative protein production (Presnyak et al., 2015).

Beyond their influence on translation, codon usage and optimality has also been shown to impact mRNA stability (Harigaya and Parker, 2017; Presnyak et al., 2015). Studies in *S. cerevisiae* have demonstrated a positive correlation between codon optimality of an mRNA and its stability (Presnyak et al., 2015; Webster et al., 2018). Triggered by slow ribosome transit, reporter mRNAs containing non-optimal codons were destabilized by Dhh1p-mediated decapping and Caf1p/Pop2p-mediated deadenylation (Radhakrishnan et al., 2016; Webster et al., 2018). Accordingly, substitution of optimal codons with non-optimal codons reduced half-lives and increased deadenylation rates for several mRNAs tested, and vice versa (Presnyak et al., 2015). This phenomenon can be extended to embryo development in *Xenopus*, *Drosophila*, and mice, in which subsets of mRNAs are more efficiently degraded based on codon optimality (Bazzini et al., 2016; Mishima and Tomari, 2016). Again, this effect is dependent on translation, but also proximity of the non-optimal codons to the 3' end.

Codon usage and optimality can also rapidly modulate the cellular gene expression program in response to stress. For example, amino acid starvation can affect translation elongation of subsets of mRNAs differently based on their codon usage (Darnell et al., 2018; Saikia et al., 2016). In HEK293 cells, amino acid deprivation downregulated translation of ribosomal protein genes – devoid of non-optimal codons – while upregulating translation of genes in the protein degradation pathway – enriched for non-optimal codons. Moreover, amino acid starvation led to the specific reduction of aminoacylation and ribosome occupation of tRNAs cognate

for optimal codons (Saikia et al., 2016). In comparison, when deprived of only a single amino acid, ribosomes did not pause at every synonymous codon for that amino acid, but only specific species (Darnell et al., 2018). Moreover, aminoacylation was reduced for the tRNA species cognate to the synonymous codons at which ribosomes paused. The physiological significance of this response is highlighted by its conservation in the bacteria *Escherichia coli* (*E. coli*) (Dittmar et al., 2005; Subramaniam et al., 2014).

Codon usage and optimality represent an evolutionarily-molded means for mRNA to control their own translation. Due to its inherent link to protein production, codon-mediated gene regulation allows a rapid adaptation to changing growth conditions. In the case of nutrient limitation, codon usage can promote the translation of essential response proteins despite an overall decrease in global protein synthesis (Saikia et al., 2016; Subramaniam et al., 2014). The complexity of this form of regulation is underscored by the finding that tRNAs charged with the same amino acid also vary in abundance during cell proliferation and differentiation (Gingold et al., 2014). Many studies now show the pervasive influence of codon usage on gene expression, which will be an interesting aspect of emerging research.

1.3 – Codon usage influences translation elongation kinetics and protein synthesis

Translation rate over a coding region is not constant and local differences in codon optimality can greatly influence elongation kinetics and co-translational folding (Rodnina, 2016). A bioinformatics analysis spanning the three domains of life showed, using genome sequence and ribosome profiling data, a widespread "ramp" of ~30-50 non-optimal codons just after translation initiation sites (Tuller et al., 2010). This ramp is translated with low efficiency, as observed by higher ribosome density (Ingolia et al., 2009), and is proposed to reduce unfavourable ribosome traffic jams at coding regions further downstream (Tuller et al., 2010).

A similar strategy is used by ER-bound proteins to facilitate their translocation by the signal recognition particle (SRP) (Pechmann et al., 2014). SRP recognizes and binds nascent peptides of ER-targeted proteins co-translationally via an N-terminal hydrophobic signal sequence (SS) or a transmembrane domain (TM). This interaction facilitates translocation of the complex to the ER where translation continues concurrently with insertion of the translating protein into the ER lumen or membrane (Akopian et al., 2013). Bioinformatic analysis revealed a region of non-optimal codons

that sits ~40 codons downstream of the SS or TM start, which the authors called an mRNA-encoded slowdown of translation (REST) element (Pechmann et al., 2014). They proposed that slow ribosome transit through the REST element promotes ER-targeting by facilitating SRP-recognition of the SS or TM as it emerges from the ribosome exit channel. Additional analysis also showed that where two TMs are separated by ~32-35 codons/amino acids, the second TM is encoded by non-optimal codons and functions as a REST element. In support, optimizing codons in the REST element of a reporter decreased ER translocation *in vivo*.

Codon-mediated local slowdowns of translation were proposed in 1987, and subsequently shown experimentally, to aid co-translational folding of nascent peptides and promote native protein folds (Purvis et al., 1987; Quax et al., 2015). Changes in protein folding would expectedly affect protein function *in vivo*. In *Neurospora*, codon-optimized versions of the circadian oscillator component *frequency* (*frq*) were unable to rescue the conidiation rhythm phenotype of *frq* null mutants (Zhou et al., 2013). Similarly, a codon-optimized version of the circadian oscillator component *Drosophila period* (*dper*) also failed to rescue the circadian clock phenotype of null *per^0* flies (Fu et al., 2016). In both these studies, function of the codon-optimized FRQ and dPER proteins were reduced compared to their wild-type versions, despite being more abundantly expressed (Fu et al., 2016; Zhou et al., 2013). This was accompanied by disruption of the timing of the proteins' turnover and phosphorylation. Notably, the impaired function of codon-optimized FRQ and dPER came as a result of non-native protein folds, as observed by altered susceptibility to Trypsin digestion. To corroborate the effect of codon-mediated translation speed on co-translational folding and protein function, slowing translation in *Neurospora* (by growth at 18°C) shifted the Trypsin susceptibility of codon-optimized FRQ closer to that of wild-type FRQ (Zhou et al., 2013).

In mammals, a silent single nucleotide polymorphism (SNP) in the multidrug resistance (*MDR1*) gene – which encodes an efflux pump – decreased the protein's sensitivity to drug inhibitors (Kimchi-Sarfaty et al., 2007). However, unlike the cases in *Neurospora* and *Drosophila*, this SNP changes the codon to an infrequent synonymous codon and decreases its optimality. While this deoptimization did not affect mRNA or protein levels, it caused *MDR1* to adopt a different fold as shown by a conformation-specific antibody as well as a Trypsin susceptibility assay. Moreover, changing the affected codon to an even more infrequent codon further decreased

inhibitor sensitivity. The effect of this SNP was shown in human and monkey cell lines. Thus, perturbation of elongation rates in either direction can have negative consequences on co-translational folding and protein function *in vivo* (Fu et al., 2016; Kimchi-Sarfaty et al., 2007; Zhou et al., 2013).

Beyond codon usage and optimality, the arrangement of codons along an mRNA can also impact elongation kinetics and protein biosynthesis (Cannarozzi et al., 2010). This is called codon autocorrelation and describes the arrangement of codons such that successive occurrences of an amino acid can be decoded by the same species of tRNA. This reduces the need to switch tRNA species and increases translation efficiency by allowing successive ribosomes on an mRNA to reuse tRNAs through substrate channeling (Figure 1.3A and Negrutskii and Deutscher, 1991). Metabolic labeling showed that codon autocorrelation is able to increase translation speed of green fluorescent protein (GFP) by up to 29% (Cannarozzi et al., 2010). Thus, codon usage and optimality are powerful ways with which translation and protein biosynthesis can be regulated in *cis*. Aside from its importance for expressing recombinant proteins in heterologous organisms, codon usage and optimality will be an important aspect to consider in protein quality control and proteostasis, as well as disease-related "silent" mutations.

1.4 – The *S. cerevisiae* RNA-binding protein Scp160p is involved in enhancing translation elongation efficiency in the context of codon arrangement

Our lab previously reported a role for the *S. cerevisiae* RBP Scp160p in boosting translation efficiency in relation to codon correlation (Hirschmann et al., 2014). In contrast to substrate channeling, Scp160p acts on tRNAs cognate for anticorrelated codons whose arrangement on an mRNA is sparse. Ribosome affinity purification and tRNA-specific RT-qPCR was employed to assay potential changes in tRNA occupancy at ribosomes upon Scp160p depletion. This analysis showed that in the absence of Scp160p, the tRNAs with the greatest drop in occupancy at the ribosome were those cognate for anticorrelated codons. Moreover, Scp160p has been shown to interact with the eukaryotic elongation factor 1A (eEF1A) (Baum et al., 2004) and ribosomal proteins located on both the A- and E-faces of ribosomes (Figure 1.3B and Gavin et al., 2006). These observations and results led to a proposed model where Scp160p promotes reuse of deacylated-tRNAs for anticorrelated codons by preventing their diffusion (Figure 1.3A and Hirschmann et al., 2014). This would be achieved by

Scp160p bridging two successive ribosomes on an mRNA and acting as a scaffold for the assembly of a tRNA recycling complex.
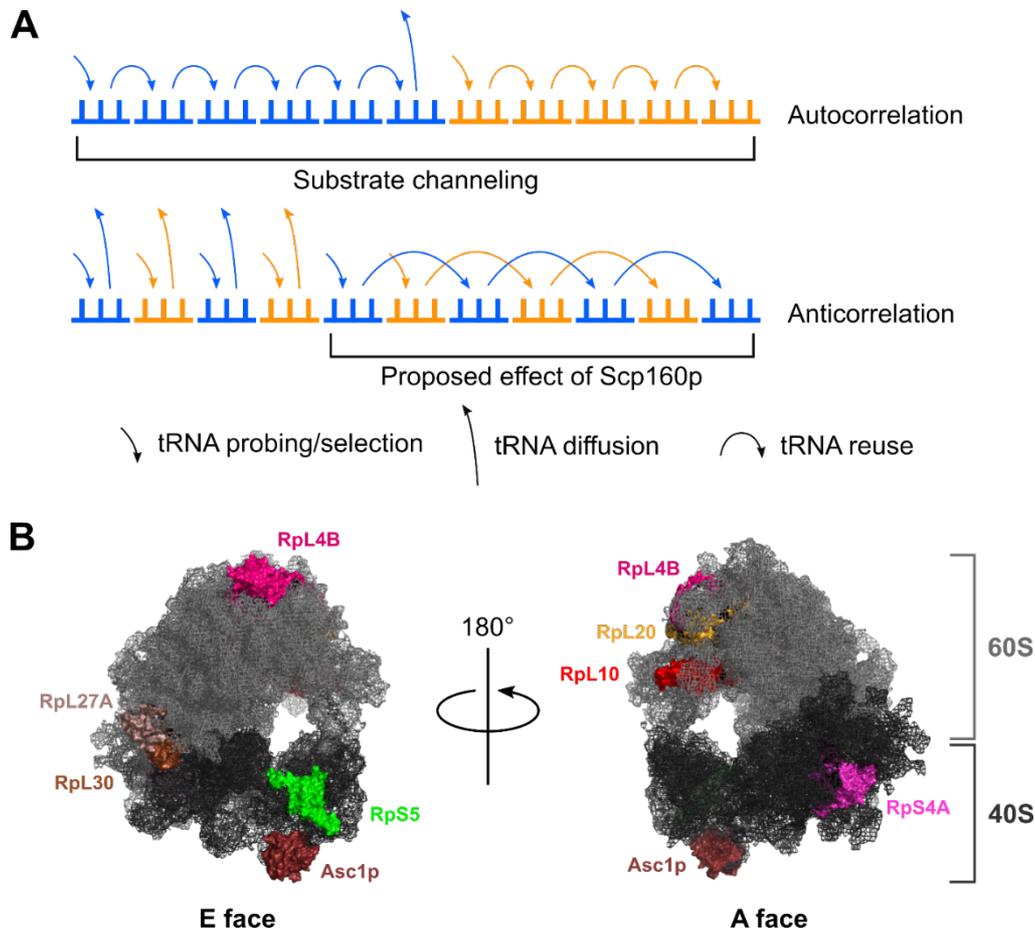
Figure 1.3 – The proposed role of Scp160p in boosting translation elongation efficiency. A) Codon autocorrelation can enhance translation efficiency by allowing tRNAs reuse through subsrate channeling. Scp160p is proposed to complement this by limiting the diffusion and/or promoting reuse of deacylated-tRNAs cognate for anticorrelated codons. Blue and yellow show synonymous codons that can use isoaccepting tRNAs. **B)** Scp160p interacts with ribosomal proteins (coloured) located on both the A- (leading) and E- (trailing) faces of the ribosome. Interaction data used are from Gavin et al., 2006 and annotated on the Saccharomyces Genome Database. *S. cerevisiae* 80S ribosome structure is from Ben-Shem et al., 2011 (PDB 4V88).

1.5 – A jack of all trades: The RNA-binding protein vigilin (Review Article: Cheng and Jansen, 2017)

Matthew H.K. Cheng and Ralf-Peter Jansen

Summary of Review Article (Full-text in Appendix):

The vigilin family of RBPs is evolutionarily conserved and is characterized by a domain architecture of 14-15 nucleic acid-binding hnRNP K-homology (KH) domains. 2017 marked 30 years since the identification of the human vigilin protein, then named high-density lipoprotein binding protein (HDLBP). Subsequently, vigilin homologs have

been identified in organisms including chicken, cow, mice, yeasts, *Xenopus*, *Drosophila*, and zebrafish. Studies in these organisms showed vigilin to be pleiotropic, with roles in many cellular processes. The subcellular distributions of the vigilin proteins is primarily cytoplasmic, with enrichment at the cytoplasmic face of the rough ER where they associate with polysomes. In higher eukaryotes, vigilin can also be detected in the nucleus associated with hetero- but not euchromatin. These observations suggest both nuclear and cytoplasmic roles for vigilin proteins.

Studies from *Drosophila* and human cell lines have linked vigilin in the nucleus to heterochromatin formation, maintenance of DNA fidelity, and gene imprinting. In contrast, studies primarily of the *S cerevisiae* vigilin homolog (Scp160p) have demonstrated roles in cytoplasmic processes such as pheromone sensing and mating; mRNA localization; as well as the regulation of P-body formation, mRNA stability and translation. The role of vigilin in the two compartments are supported by the protein's putative nuclear localization and nuclear export signals. Indeed, the human vigilin can shuttle between the nucleus and cytoplasm as part of a tRNA-shuttling complex.

While studies in different organisms show vigilin's widespread influence of cellular processes, they also provide a disparate view of the proteins' functions. In the review, we presented a brief historical account of vigilin research and a concerted summary of the major findings, allowing an easier comparison of the protein's various roles in cell function.

Addendum to Review Article:

During the preparation of this review, two studies reported additional roles for vigilin in the regulation of stress-induced ligand expression in the human innate immune system (Berhani et al., 2017) and in *Caenorhabditis elegans* larval development (Zabinsky et al., 2017). A third study was published shortly after the review, which expanded upon the role of vigilin in gene imprinting (Yu et al., 2018). The major findings of these studies are summarized in this subsection, to be considered in addition to the review (see Appendix, Cheng and Jansen, 2017).

In the study from Berhani and colleagues (2017), vigilin was found to regulate expression of a factor involved in stress-induced activation of natural killer (NK) cells in the innate immune system. NK cells are activated during stresses such as viral infection, by cell-surface receptors binding of stress-induced ligands (Raulet et al., 2013). Vigilin was found to bind the 5'-UTR of the mRNA encoding stress-induced

ligand MHC Class I-related chain B (MICB) and downregulate protein expression at the level of translation but not stability (Berhani et al., 2017). Accordingly, shRNA-mediated knockdown of vigilin increased NK cell activation. Although regulation of MICB expression has been reported to involve many miRNAs and RBP acting on the 3'-UTR (Raulet et al., 2013), this study uncovered a novel means of MICB regulation by vigilin via its 5'-UTR (Berhani et al., 2017). However, it remains unclear what mechanism(s) vigilin employs to inhibit MCIB translation. Identification of additional proteins which bind or interact with vigilin at the MICB 5'-UTR will provide further insight into this question. Furthermore, given the many factors acting through the MICB 3'-UTR, it will be important to study how they coordinate regulation with vigilin.

The study from Zabinsky and colleagues (2017) looked for mutations which could enhance an L1 larval development arrest phenotype associated with the loss of ALG1-interacting protein 2 (*AIN-2*) function in *C. elegans*. *AIN-2* is one of two GW182 homologs. The authors found that vigilin loss-of-function (or depletion by RNAi) enhanced this phenotype in *ain-2* mutants, linking vigilin to the miRISC machinery. In support, vigilin loss-of-function was also synergistic with deletion of several miRNA families involved in development. Notably, seeding sites for these miRNA families were enriched in hundreds of mRNAs that co-immunoprecipitated with GFP-tagged vigilin. The relationship between vigilin and the miRISC, although interesting, remain unclear. The authors reasoned it to be unlikely that vigilin acts as a co-repressor of *AIN-2* and the miRISC based on the following observations: 1) They did not observe a physical interaction between vigilin and *AIN-2* or Dicer-2, and 2) The synergistic effect of vigilin loss-of-function with deletions of various miRNA families suggest they function in parallel pathways (Zabinsky et al., 2017). Indeed, vigilin's role in larval development may be complementary to that of miRISC and function via another form of mRNA regulation like localization. This may be teased apart by analyzing the translation, stability, and/or localization of several candidate vigilin mRNAs targets identified in this study.

The third study, from Yu and colleagues (2018), expands upon their previous work on the function of vigilin in gene imprinting. Gene imprinting is an epigenetic form of gene regulation that results in monoallelic expression from only one parental allele (Macdonald, 2012). As outlined in my review, vigilin was found to bind the regulatory regions of two imprinted genes, insulin-like growth factor 2 (*Igf2*) and *H19*, in a manner dependent on the CCCTC-binding factor (CTCF) (Cheng and Jansen, 2017; Liu et al.,

2014b; Shen et al., 2014). Yu and colleagues now report that knockdown of vigilin or CTCF result in the specific upregulation of *Igf2* from the imprinted (normally silenced) allele (Yu et al., 2018). Furthermore, they report that vigilin domains KH1-7 interacts with the zinc-finger domain of CTCF *in vitro* and that this interaction *in vivo* is RNA-dependent. However, RNase A treatment did not completely abolish this interaction and they were unable to identify the RNA responsible by RNA immunoprecipitation. Nevertheless, the authors reasoned and observed upon shRNA knockdown that the vigilin-CTCF interaction is facilitated by the lncRNA H19, whose expression they previously shown to be suppressed by vigilin and CTCF (Liu et al., 2014b). It remains to be tested whether the vigilin-CTCF-lncRNA H19 complex includes other factors and whether the vigilin domains KH1-7 interact with CTCF *in vivo*. It is also unknown whether vigilin binds lncRNA H19 and if its knockdown affects the vigilin-CTCF complex directly.

The studies highlighted here and in Cheng and Jansen (2017) show the multifaceted and expanding role of the vigilin proteins. Yet our knowledge of the functional mechanisms of this family of RBPs function remain unclear. Despite the evidence that vigilin shows specificity for certain RNAs (Berhani et al., 2017; Cunningham et al., 2000; Li et al., 2003; Mobin et al., 2016), there lacks a consensus binding-motif. This, and the ability of vigilin to bind many classes of RNAs, put into question how target RNAs are recognized. Furthermore, given vigilin proteins' abundance of KH domains, it would be interesting to investigate the stoichiometry with RNAs. Since vigilin associates with a number of different proteins and mRNP complexes (Frey et al., 2001; Kruse et al., 2000; Lang and Fridovich-Keil, 2000; Liu et al., 2014b; Lu et al., 2012; Sezen et al., 2009; Wang et al., 2005), it is possible that RNA specificity is dictated by interaction partners. In depth structure-function analyses of vigilin will shed light on how RNA-binding and regulation is coordinated.

1.6 – Polyglutamine and low-complexity regions in mediating protein aggregation and phase separation

This work made use of reporters containing a 97-104 residue long glutamine repeat (a polyglutamine (polyQ) region), a well-studied example of low-complexity regions (LCRs). LCRs are so-called due to their low diversity in overall amino acid make-up and are enriched for amino acids like the polar glutamine (Q), asparagine (N), and serine (S); the positively charged lysine (K) and arginine (R); the negatively

charged glutamic acid (E) and aspartic acid (D); the aromatic phenylalanine (F) and tyrosine (Y), as well as the non-polar glycine (G) (Brangwynne et al., 2015). These amino acids are commonly arranged in tandem repeats in LCRs, forming motifs that include F/G- and Q/N-rich regions (Bergeron-Sandoval et al., 2016; Brangwynne et al., 2015). These properties of LCRs predisposes the parent proteins to aberrant aggregation and amyloid formation and are responsible for a number of diseases.

1.6.1 – An expanded polyQ region in Huntington's disease

Expanded polyQ regions are associated with a number of neurodegenerative diseases collectively called repeat expansion disorders. They include Huntington's disease (HD), spinobulbar muscular atrophy, dentatorubral-pallidoylusian atrophy, and six variants of spinocerebellar ataxias (type 1, 2, 3, 6, 7, and 17) (Fan et al., 2014; Hoffner and Djian, 2014; Scherzinger et al., 1999). Notably, the proteins linked with these diseases only become pathogenic when their polyQ regions expand from ~20 residues to greater than a threshold of ~35. The polyQ reporters used in this work are derived from exon 1 of the expanded form of huntingtin protein (mHtt; Htt for the wild-type unexpanded form) (Fan et al., 2014).

Molecular studies of mHtt have shown the protein's propensity to aggregate spontaneously into insoluble inclusions (Hoffner and Djian, 2014; Scherzinger et al., 1999). Thus, pathogenesis comes in part from sequestration of the protein and loss of normal Htt functions (Hofer et al., 2018; Schulte and Littleton, 2011). Conversely, the aggregated form of mHtt disrupts various cellular processes, resulting in toxic gain of functions (Fan et al., 2014; Hofer et al., 2018; Hoffner and Djian, 2014). For example, cytoplasmic aggregates of a mHtt reporter were observed to impair nucleocytoplasmic transport (Woerner et al., 2016). In mammalian cell lines and mice brains, aggregated mHtt caused aberrant mislocalization of the transport factors THOC2, importin α-1, and importin α-3 to the cytoplasm. Moreover, mHtt aggregates also caused mislocalization and aggregation of nuclear pore complex components (Grima et al., 2017). Finally, mHtt aggregates in *S. cerevisiae* and HeLa cells can sequester the HSP40 chaperones Sis1p (in *S. cerevisiae*) and DnaJB1 (in mammals), thereby blocking their function in the transport of misfolded client proteins to the nucleus for degradation (Park et al., 2013).

In addition to nucleocytoplasmic transport factors, membrane integrity is also disrupted by aggregated expanded polyQ proteins. Aggregated and fibrillary forms of

synthetic 35Q peptides caused leakage – but not fragmentation – of unilamellar vesicles (Ho et al., 2016). Further investigation by atomic force microscopy showed that these fibrils disrupted and increased rigidity of supported lipid bilayers. In agreement, correlative light and electron microscopy (CLEM) revealed extensive interaction between mHtt-based polyQ inclusions and the ER *in situ* (Bäuerlein et al., 2017). This caused impaired ER dynamics and ER breakage. Perhaps not surprisingly given the ER damage, mHtt can also impair ER-associated degradation (ERAD) and the unfolded protein response (UPR) in *S. cerevisiae* and mammalian neuron-like cells (Duennwald and Lindquist, 2008). However, the impairment of these pathways may not be membrane-related and is attributed to the entrapment of the ERAD factors Cdc48p/p97, Npl4p, and Ufd1p with mHtt inclusions.

There is also evidence that cellular toxicity is conferred by the soluble oligomeric form of mHtt instead of the insoluble inclusions (Kim et al., 2016; Leitman et al., 2013). Leitman et al. (2013) reported that in several mammalian cell lines, ER-stress is triggered concomitantly with mHtt oligomerization, and precedes the appearance of visible insoluble inclusions. Moreover, this is accompanied by impairment of ERAD and accumulation of ERAD substrates. While this seems to be in disagreement from the report from Duennwald and Lindquist (2008), this discrepancy likely results from the different substrates observed in the two studies (ERAD factors vs substrate). Therefore, it is possible that ERAD impairment begins with soluble mHtt and is maintained by subsequent sequestration of important ERAD factors by insoluble mHtt inclusions.

Further support for soluble mHtt oligomers as the toxic species comes from a study of their interactome in mice neuroblastoma cells. The interactome of soluble mHtt oligomers was enriched for proteins involved in a variety of cellular processes including RNA binding/translation, ribosome biogenesis, DNA binding/transcription, intracellular transport, and chaperone function (Kim et al., 2016). These interactions were mainly mediated by LCRs. The interaction of these key cellular proteins was specific to soluble mHtt oligomers as they are largely excluded from the interactome of insoluble mHtt inclusions. Instead, the interactome of insoluble mHtt inclusions was enriched for chaperones, which prompted the authors to propose that bulk toxicity is conferred by soluble mHtt oligomers and that mHtt inclusions are cytoprotective. In support, mHtt is targeted to the insoluble protein deposits (IPOD), a compartment for protein quality control and degradation of terminally misfolded proteins (Hofer et al.,

14

2018; Kaganovich et al., 2008). Together, these studies demonstrate the adverse effects of mHtt on many cellular processes and structures.

1.6.2 – Low-complexity regions in protein function

Another class of LCR are the prion-like domains (PrDs), which are generally Q/N-rich and defined based on similarity to the aggregation-mediating domains of *S. cerevisiae* prions (Li et al., 2018). Prions are the inheritable self-propagating isoforms which certain proteins can spontaneously convert to. One of the best studied prions is [PSI+], the converted form of the translation terminator Sup35p in *S. cerevisiae* (Glover et al., 1997; Liebman and Chernoff, 2012). In humans, mutations in PrD-containing proteins that cause their aberrant folding, aggregation, and function have been linked to neurodegenerative diseases including amyotrophic lateral sclerosis, frontotemporal lobar degeneration, Creutzfeldt-Jakob disease, and Alzheimer's disease (King et al., 2012). A commonality between expanded polyQ and Q/N-rich regions is their ability to facilitate their encompassing proteins to adopt an anti-parallel ß-sheet fold, nucleate and self-assemble into amyloid fibrils (Hoffner and Djian, 2014; Liebman and Chernoff, 2012).

Despite their potential detriment, polyQ and Q/N-rich regions are common features of the eukaryotic proteome (Alexandrov and Ter-Avanesyan, 2013; King et al., 2012; Michelitsch and Weissman, 2000; Schaefer et al., 2012). Sequence analyses estimate that ~1 and 1.69% of the human and *S. cerevisiae* proteomes, respectively, contain Q/N-rich regions (King et al., 2012; Michelitsch and Weissman, 2000). In an extreme example, ~14% of the proteome of the slime mold *Dictyostelium discoideum* is predicted to contain Q/N-rich regions (Malinovska et al., 2015). This prevalence in eukaryotic proteomes hints at a biological importance of these regions. Indeed, polyQ-containing proteins are enriched among human protein complexes and the length of the polyQ regions correlate with the number of interaction partners (Schaefer et al., 2012). Likewise, prion/prionogenic proteins have a tendency to interact with other Q/N-rich proteins in *S. cerevisiae* (Harbi and Harrison, 2014). Thus, in their native unexpanded context, these LCRs may be evolutionarily conserved effectors of protein-protein interaction.

Recently, many studies have demonstrated a role for polyQ, Q/N-rich regions, and other LCRs in phase separation. Phase separation is a phenomenon in which proteins and nucleic acid de-mix from the cellular milieu into viscous supersaturated

liquid-like droplets or membrane-less organelles (Brangwynne et al., 2015). This process is driven by intra- and intermolecular electrostatic interactions of the proteins/side chains and nucleic acids involved. As such, it can be modulated by post-translational modifications as well as the ionic properties of the environment (Bergeron-Sandoval et al., 2016; Brangwynne et al., 2015). Phase separation has been shown to drive the formation of mRNPs like P-bodies, stress granules (SG), cellular bodies in oocytes and embryos, as well as modulate protein function in regular cellular processes (Boke et al., 2016; Decker et al., 2007; Molliex et al., 2015; Rabouille and Alberti, 2017; Reijns et al., 2008).

Under conditions of stress, *S. cerevisiae* Lsm4p (a subunit of the Lsm1-7p complex) is recruited to P-bodies via its Q/N-rich region (Decker et al., 2007; Reijns et al., 2008). Similarly, the pyruvate kinase Cdc19p undergoes LCR-mediated amyloid formation and localization into SGs during certain stresses and in stationary phase (Saad et al., 2017). Finally, Sup35p phase separates into condensates in *S. cerevisiae* and *Schizosaccharomyces pombe* (*S. pombe*) depleted of energy, in a process mediated by the protein's PrD (Franzmann et al., 2018). The importance of LCR-mediated stress response has also been shown in mammalian cells. Recruitment of SG components hnRNPA1 and TIA-1 to SGs is dependent on the proteins' LCR and Q-rich regions, respectively (Gilks et al., 2004; Molliex et al., 2015). Strikingly, the Q-rich region of TIA-1 can be functionally replaced with the PrD of the *S. cerevisiae* Sup35p, highlighting a conservation of function for these regions (Gilks et al., 2004).

LCR-mediated protein aggregation is also important outside of stress. PolyQ regions have also been shown to promote transcription activity in *S. cerevisiae*, *Drosophila*, and mammalian cells (Atanesyan et al., 2012). In many cancers, the LCRs of the RBPs FUS, EWS, and TAF15 are translocated to various DNA-binding domains. Remarkably the LCRs functioned as transcription activator domains in the resulting fusion proteins (Kwon et al., 2013). This required the LCR-mediated formation of amyloid fibers, which can then interact with the un-phosphorylated C-terminal domain of RNA polymerase II.

In the filamentous multinucleate fungi *Ashbya gossypii*, the polyQ region of the RBP Whi3p is required for its ability to cluster and localize target mRNAs by forming phase-separated mRNPs (Lee et al., 2013, 2015). Whi3p localizes the cyclin-encoding *CLN3* mRNA to distinct nuclei and drives asynchronous nuclear division within a syncytium (Lee et al., 2013). Alternatively, Whi3p localizes the formin-encoding *BNI1*

and polarisome-encoding *SPA2* mRNAs to sites of polarized growth and facilitates asymmetry breaking (Lee et al., 2015). Remarkably, *CLN3*/Whi3p and *BNI1*/*SPA2*/Whi3p droplets are prevented from mixing with each other based on the inter-RNA interactions and secondary structures of the cargo mRNAs (Langdon et al., 2018).

Finally, the *Drosophila* cytoplasmic polyadenylation element binding protein family member Orb2 can functionally as a translation repressor or activator based on its amyloid status (Khan et al., 2015). Amyloid formation of Orb2 is dependent on its N-terminal Q-rich PrD. Monomeric Orb2 acts as a translation repressor that promotes deadenylation and decay of target transcripts while amyloidogenic Orb2 acts as a translation activator that stabilizes poly(A) tails and target transcripts. As Orb2 regulates translation at the synapse this distinction has important ramifications for long-term memory formation.

A wealth of research has shown that the dynamics of these protein aggregates/inclusions lie on a spectrum ranging from liquid-like to solid-like amyloids. Indeed, LCR aggregates have been observed to transition from liquid- to solid-like states over time *in vitro* and *in vivo* (Lin et al., 2015; Patel et al., 2015; Peskett et al., 2018; Zhang et al., 2015). Importantly, these transitions were accompanied by a conformational change of these protein aggregates into amyloid fibers (Lin et al., 2015; Patel et al., 2015; Peskett et al., 2018; Zhang et al., 2015). Thus, the biological functions of LCR-containing proteins lie on a fine balance between their monomeric forms and insoluble amyloid formation. The tipping of this system too far in either direction can lead to detrimental effects and disease.

2. AIMS OF THE PRESENT WORK

Our lab has previously reported a role for the *S. cerevisiae* vigilin homolog, Scp160p, in enhancing translation elongation efficiency of a subset of mRNAs in the context of codon arrangement (see section 1.4 and Hirschmann et al., 2014)). The aim of this work is to further investigate this function of Scp160p and how it might influence protein biosynthesis and folding. As local changes in elongation speed mediated by codon usage can impact protein folding/aggregation (see section 1.3 and Jacobson and Clark, 2016; Pechmann et al., 2014; Yu et al., 2015), I employed three polyQ reporters differing in the codon composition of their polyQ-encoding regions (Figure 2) and assessed how loss of Scp160p influences the aggregation of these reporters. Given the prevalence of polyQ and Q/N-rich regions in eukaryotic proteomes, I also probed the potential physiological impact of Scp160p on aggregation of the *S. cerevisiae* proteome using an approach that combined filter-trap binding and quantitative mass spectrometry. An on-going question is via what mechanism(s) Scp160p influences polyQ- and Q/N-mediated protein aggregation. An additional part of my project investigated the functional relationship between Scp160p and Bfr1p in ploidy maintenance and expression of the polyQ reporter proteins.
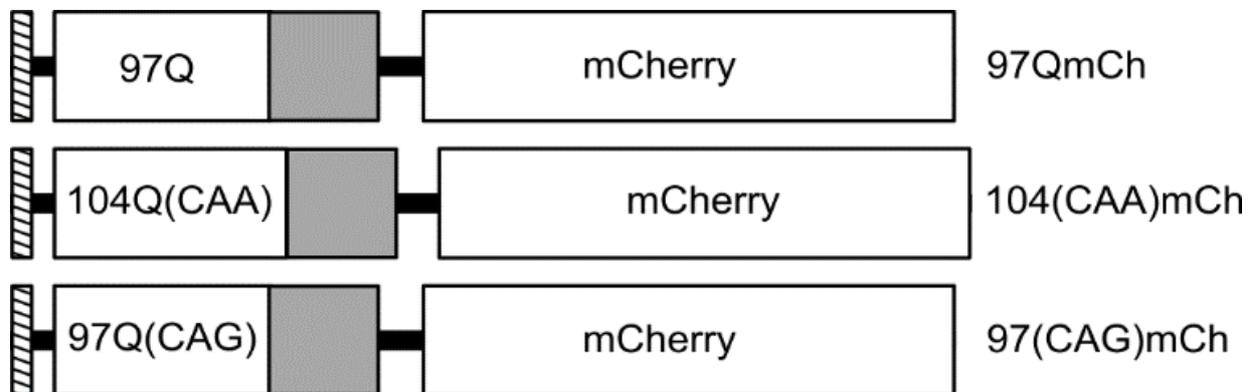
Figure 2 – Schematics of the polyQ reporters used in this study. The reporters used in this study are based on exon 1 of the expanded Huntingtin (mHtt) including the proline-rich region (grey). The reporters differ in their codon usage in the polyQ-encoding region. In *S. cerevisiae*, CAA is the frequent codon and CAG is the infrequent codon. Note that only the 97Q(CAG)mCh reporter accurately represents mHtt at both the nucleic-acid and protein levels. All three reporters have an N-terminal myc epitope (hatched) and a C-terminal mCherry tag.

3. RESULTS

3.1 – Loss of the RNA-binding proteins Scp160p and Bfr1p result in increased ploidy in *S. cerevisiae*

In *S. cerevisiae*, Scp160p has been shown to interact and function closely with another RBP Bfr1p. Bfr1p is involved in nuclear segregation and the early secretory pathway (Jackson and Képès, 1994). It was identified through a multi-copy suppressor screen to confer partial resistance to the drug brefeldin A – which acts on the secretory pathway – in a sensitized mutant. Interestingly, although Bfr1p has been shown to associate with RNAs (Hogan et al., 2008; Kramer et al., 2014; Lapointe et al., 2015), structural prediction shows that the protein has no conventional RBDs but instead contains three coiled-coil domains (Jackson and Képès, 1994).

Unlike Scp160p, Bfr1p is only conserved among the *Ascomycota* phylum (unpublished, Srinivas Manchalu), arguing that the relationship between the two RBPs is fungal-specific. Nevertheless, several lines of evidence suggest that Scp160p and Bfr1p have a close functional relationship. An immunoprecipitation and microarray approach found that both proteins are each associated with >1,000 mRNAs, a large fraction of which are common targets (Hogan et al., 2008). *in vivo* studies have also shown that Scp160p and Bfr1p have similar subcellular distributions, which are mainly cytoplasmic with enrichment at the ER (Lang et al., 2001; Weidner et al., 2014) where they interact with P-body components and regulate the formation of P-bodies (Weidner et al., 2014). Moreover, Scp160p association with polysomes is dependent on Bfr1p and the two RBPs belong to the same polysome-associating mRNP that also includes the poly(A)-binding protein Pab1p (Lang and Fridovich-Keil, 2000; Lang et al., 2001). Finally, a striking phenotype shared by the *scp160Δ* and *bfr1Δ* cells is an increase in ploidy or DNA content (Jackson and Képès, 1994; Weber et al., 1997; Wintersberger et al., 1995). Although it is unclear whether the two function together in ploidy maintenance, one study has shown Scp160p is part of a regulatory complex called SESA (Smy2p-Eap1p-Asc1p-Scp160p) which provides a potential link to ploidy maintenance (see Appendix, Cheng and Jansen, 2017). SESA downregulates translation of *POM34*, a component of the nuclear pore complex (NPC), in response to defects in spindle pole body (SPB) duplication (Sezen et al., 2009). This study also reported a possible but untested link between Bfr1p and the SESA network.

3.1.1 – scp160Δ *and* bfr1Δ *do not show additive effect in ploidy phenotype*

To better understand how Scp160p and Bfr1p contribute to ploidy maintenance, and whether they function together in this regard, I generated an *scp160Δbfr1Δ* double deletion strain and assessed DNA content by propidium iodide (PI) staining and flow cytometry. As deletion of either *SCP160* or *BFR1* leads to ploidy increase, subsequent deletions of these two genes by standard PCR-based method was not feasible (Jackson and Képès, 1994; Longtine et al., 1998; Wintersberger et al., 1995). Therefore, the double deletion mutant was generated first by creating a heterozygous *scp160Δ/SCP160 bfr1Δ/BFR1* diploid strain. Sporulation was induced in the resulting heterozygous diploid and the *scp160Δbfr1Δ* haploid was recovered by tetrad dissection and marker selection. The double deletion and resulting mating type were confirmed by genomic PCR. Figure 3.1.1 shows that the ploidy of *scp160Δbfr1Δ* cells is increased to a similar extent as the single mutants. This result suggests that Scp160p and Bfr1p likely contribute to ploidy maintenance through a common or redundant mechanism(s). Alternatively, it may be possible that the deletions affect different aspects of ploidy maintenance that nonetheless together result only in a doubling in ploidy.
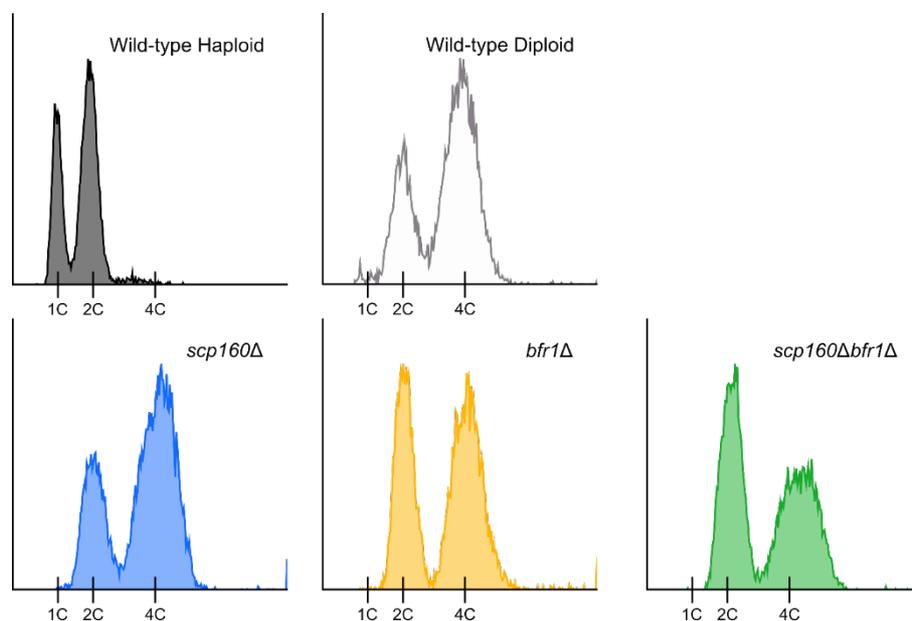
Figure 3.1.1 – *scp160Δ* and *bfr1Δ* are not synergistic. Propidium iodide staining and flow cytometry show that *scp160Δ* and *bfr1Δ* cells have diploid DNA contents. *scp160Δbfr1Δ* also show a diploid DNA content suggesting the two deletions are not synergistic.

3.1.2 – KH11-14 are not required for ploidy maintenance

Previous studies have shown that loss of Scp160p's KH14 domain reduces polysome association (Li et al., 2004) and truncating KH13 and 14 abolishes binding of Scp160p to target mRNAs (Hirschmann et al., 2014). These observations demonstrate the importance of the C-terminal KH domains. To investigate whether these and other KH domains might also contribute to ploidy maintenance, I generated a series of truncations removing successive KH domains from the C-terminus of Scp160p. A GFP tag was appended to the C-terminus of the truncated proteins allowing their detection. Western blotting shows that all truncations are stable, although present with variations in abundance.

Flow cytometry of PI-stained cells showed that truncations up to and including KH11 had ploidy similar to that observed in wild-type haploid cells (Figure 3.1.2), arguing that KH11-14 are not necessary for maintenance of exact ploidy. This contrasts with previous observations, in which KH11 and/or 12 were necessary for maintaining a haploid DNA content (Baum et al., 2004). This discrepancy may lie in the different *S. cerevisiae* background strains used in the studies (S288C in Baum et al., 2004 vs W303 for this study) and method of construct generation. Nevertheless, in the W303 background, KH11-14 are not required for ploidy maintenance.

Strikingly, although the Scp160p(KH1-10)-GFP truncation (in which KH11-14 are lost) remains haploid, the Scp160p(KH1-11)-GFP truncation (in which only KH12-14 are lost) shows a diploid profile. Western blot shows that all truncations are expressed, albeit at various levels. Additionally, the comparable levels of Scp160p(KH1-11)-GFP in comparison to the full-length Scp160p-GFP shows that the increased ploidy for the truncation is unlikely due to aberrant folding and/or synthesis. Alternatively, while a relationship between Scp160p dosage and ploidy maintenance has not been established, the loss of KH12 may sensitize the cells' ploidy maintenance capabilities to levels of Scp160p. Such a scenario would reconcile the haploid DNA content of the Scp160p(KH1-10)-GFP truncation, in which the protein is more abundant, with the diploid DNA content of the Scp160p(KH1-11)-GFP truncation.
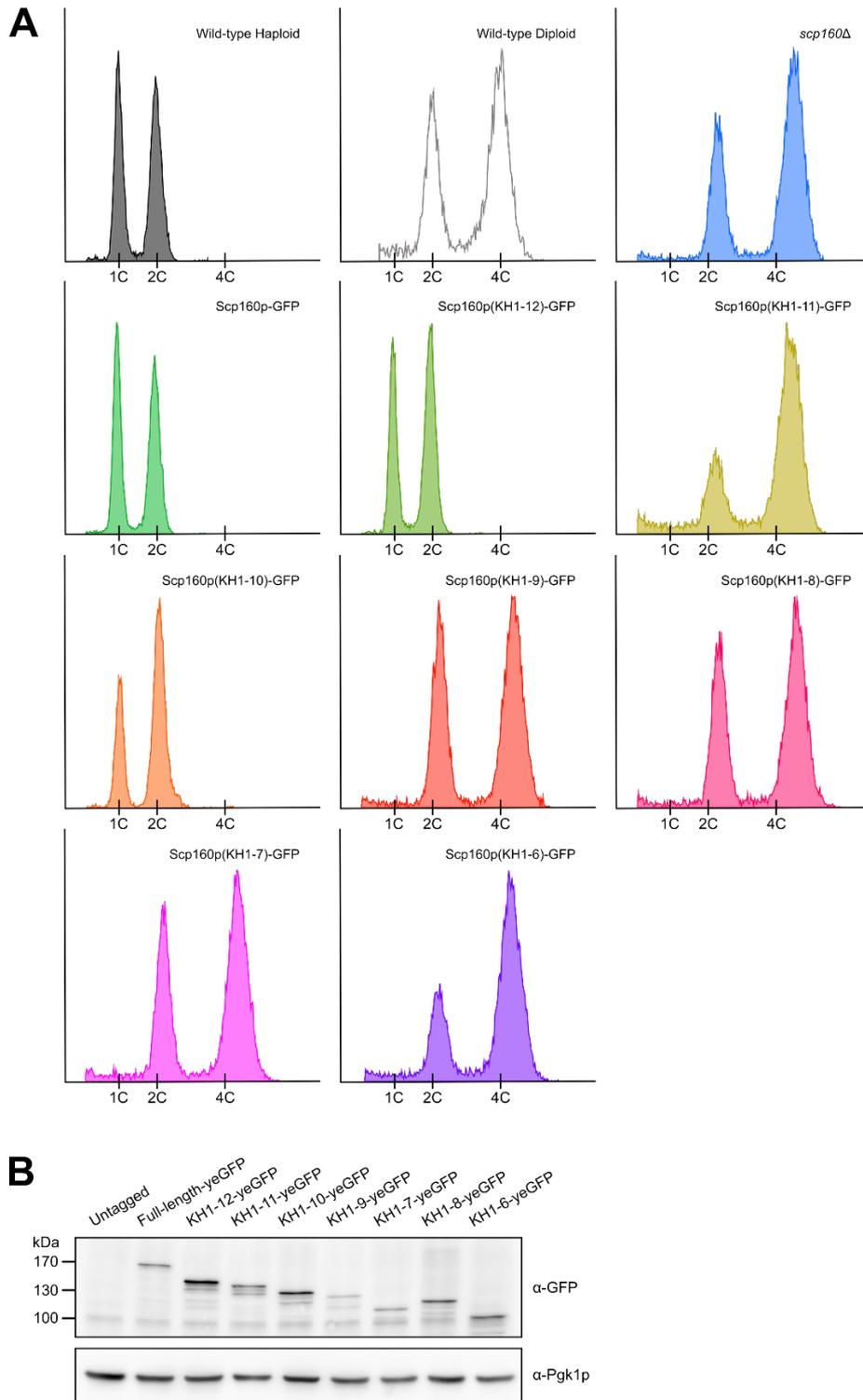
Figure 3.1.2 – KH10-14 is not required for ploidy maintenance. A) A series of Scp160p C-terminal truncations was generated and assessed for ploidy. Cells with Scp160p truncations up to and including KH11 showed haploid DNA contents. Of note is the diploid DNA content displayed by the Scp160p(KH1-11)-GFP truncation despite the haploid DNA content of the subsequent Scp160p(KH1-10) truncation. **B)** Western blot against the C-terminal GFP tag of the truncation series show that the truncations are expressed and stable, albeit with various abundances.

3.2 – The RNA-binding protein Scp160p facilitates aggregation of many Q/N-rich proteins (Research Article: Cheng et al., 2018)

Matthew H.K. Cheng, Patrick C. Hoffmann, Mirita Franz-Wachtel, Carola Sparn, Charlotte Seng, Boris Maček, Ralf-Peter Jansen

Summary of Research Article (Full-text in Appendix):

The RNA-binding protein Scp160p is the yeast homolog of the conserved family of vigilin proteins. In studies spanning multiple organisms, vigilin proteins have been associated with ploidy maintenance, chromatin remodeling, nucleocytomplasmic transport of tRNAs, regulation of P-body formation, localization of mRNAs, stability of mRNAs, and translation of mRNAs (Cortés et al., 1999; Cunningham et al., 2000; Gelin-Licht et al., 2012; Graham and Oram, 1987; Marsellach et al., 2006; McKnight et al., 1992; Mobin et al., 2016; Weidner et al., 2014; Wen et al., 2010; Wintersberger et al., 1995). Our lab has previously reported that Scp160p benefits translation of a subset of *S. cerevisiae* mRNAs via recycling of tRNAs cognate to anticorrelated codons (see section 1.4 and Hirschmann et al., 2014). In this study, we investigated this function further by assessing how loss of Scp160p might influence biosynthesis of three polyQ reporters (97QmCh, 97Q(CAG)mCh, and 104Q(CAA)mCh; Figure 2) which differ in the codon usage of their polyQ-encoding region.

We first assessed aggregation of the polyQ reporters by microscopy after inducing their expression for 3 and 7 hours and observing the fluorescent signal from their C-terminal mCherry tags. Based on the distribution of their mCherry signal, cells observed were categorized as showing 1) No Signal, 2) Soluble, 3) Small Aggregates, and 4) Large Aggregate(s) (Figure 3.2.1A and Cheng et al., 2018). Taking the percentage of cells in each category as an approximation of polyQ aggregation, we saw that aggregation was reduced in *scp160Δ* cells (Figure 3.2.1B-D). This was supported biochemically by filter-trap binding (Figure 3.2.1E-G), whereby aggregated polyQ proteins can be trapped on a nitrocellulose based on their resistance to denaturation by mild SDS treatment. Notably, loss of Scp160p resulted in the reduction of polyQ aggregation regardless of codon usage, suggesting the RBP's role here is separate from its role in tRNA reuse during elongation.
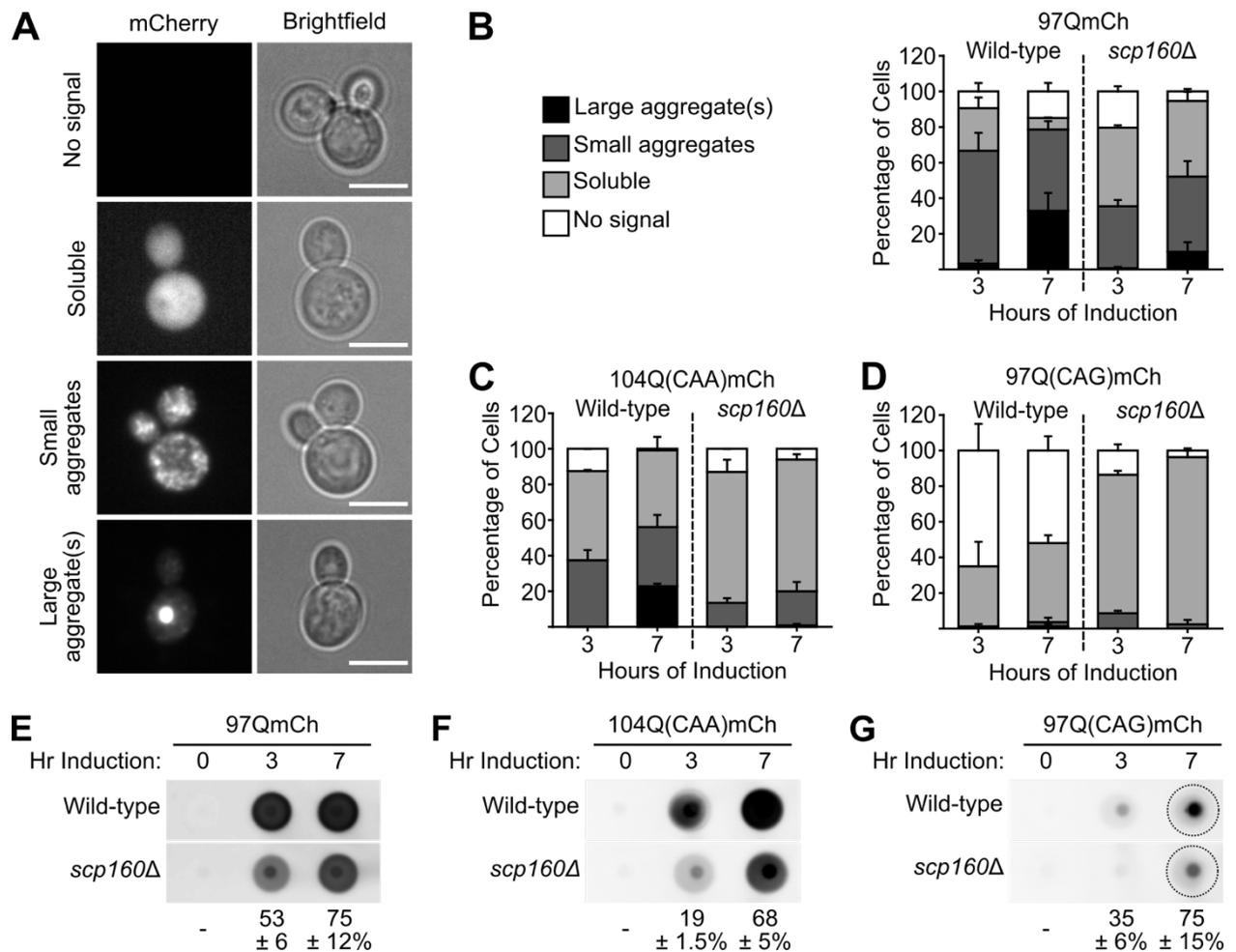
Figure 3.2.1 – Aggregation of polyQ reporter proteins is reduced in *scp160Δ* cells. Three polyQ reporters differing in the codon usage of their polyQ region were induced in wild-type and *scp160Δ* cells. **A)** The cells were were observed by microscopy and categorized based on the distribution of the reporters' mCherry signal. The percentage of cells containing aggregated **B)** 97QmCh, and **C)** 104Q(CAA)mCh were reduced in the *scp160Δ* strain. **D)** 97Q(CAG)mCh expression in wild-type cells were low, making microscopic analysis difficult. Aggregation of **E)** 97QmCh, **F)** 104Q(CAA)mCh, and **G)** 97Q(CAG)mCh were assessed by filter-trap binding. The numbers show the average immunosignal for myc of *scp160Δ* cells compared to wild-type cells. SEM is shown for three biological replicates.

In contrast, toxicity of the three polyQ reporters is codon usage-dependent with the 97QmCh and 104Q(CAA)mCh reporters conferring modest toxicity and the 97Q(CAG)mCh reporter conferring severe toxicity to wild-type cells (Figure 3.2.2A). Strikingly, 97Q(CAG)mCh toxicity is suppressed in the absence of Scp160p. Furthermore, removal of in-frame start codons of the toxic 97Q(CAG)mCh reporter abolished toxicity in wild-type cells (Figure 3.2.2B). Together, this demonstrates that codon usage, and not amino acid composition, is a major effector of polyQ-induced toxicity, and that Scp160p and translation initiation are both required for toxicity of the 97Q(CAG)mCh reporter.
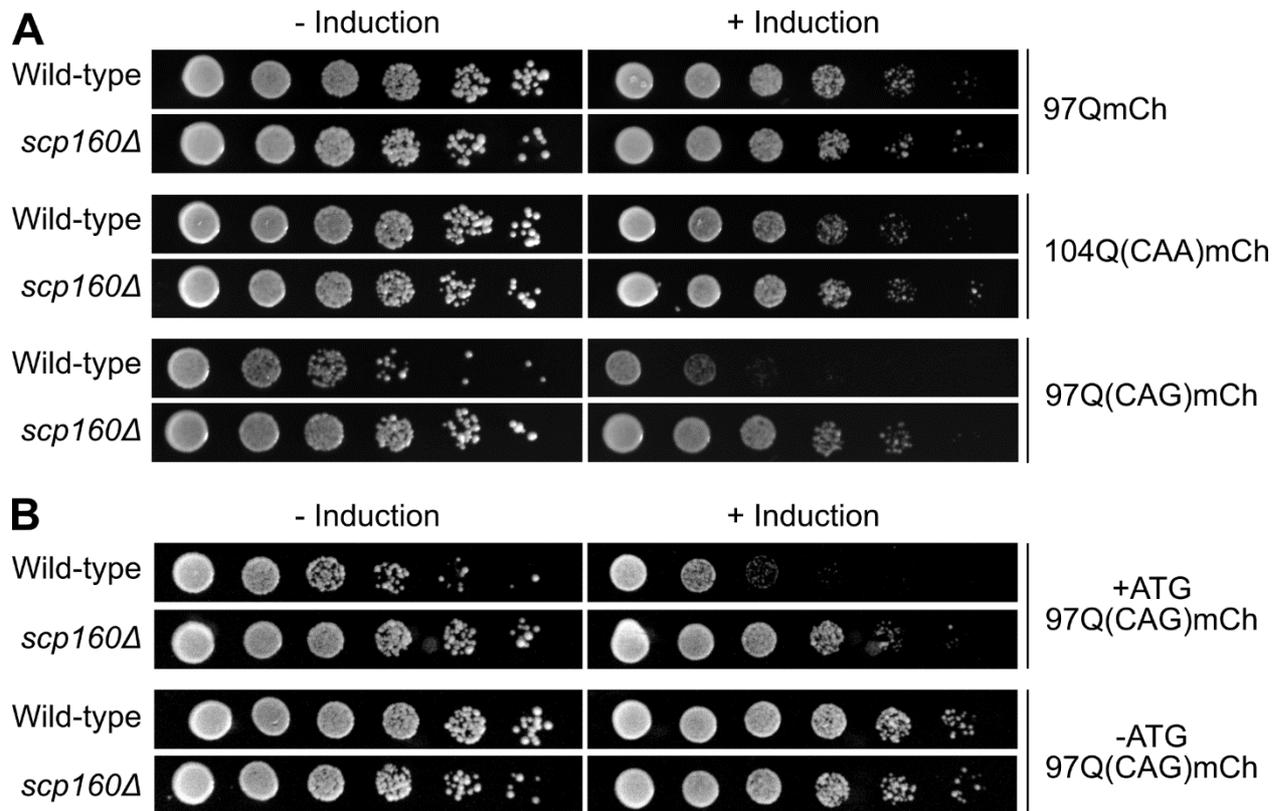
Figure 3.2.2 – Codon usage is a major determinant in polyQ-induced toxicity. A) Toxicity of 97QmCh, 104Q(CAA)mCh, and 97Q(CAG)mCh were assayed by serial dilution assay. 97QmCh and 104Q(CAA)mCh conferred modest toxicity to wild-type and *scp160Δ* cells. 97Q(CAG)mCh was severely toxic to wild-type cells and this toxicity was suppressed in *scp160Δ* cells. **B)** Mutation of the start codons upstream of the polyCAG repeats in 97Q(CAG)mCh abolished toxicity to wild-type cells.

To assess the physiological impact of Scp160p's effect on polyQ-mediated protein aggregation, we combined filter-trap binding with quantitative mass spectrometry to probe how loss of Scp160p affects the endogenous aggregated proteome in *S. cerevisiae* under normal growth conditions. We observed that aggregation of 77 proteins were altered in *scp160Δ* cells, with 48/77 showing reduced aggregation (Figure 3.2.3). Of these 48 proteins, 27 proteins (56.3% of the proteins with reduced aggregation) contained Q/N-rich regions or prion-like domains as predicted by previous publications (Alberti et al., 2009; Michelitsch and Weissman, 2000). This represents a 30-fold enrichment for proteins with Q/N-rich regions or prion-like domains over that in the yeast proteome. Among these proteins are the well-characterized prions [PSI+] (Sup35p), [OCT+] (Cyc8p), [NU+] (New1p) (Glover et al., 1997; Liebman and Chernoff, 2012; Patel et al., 2009).
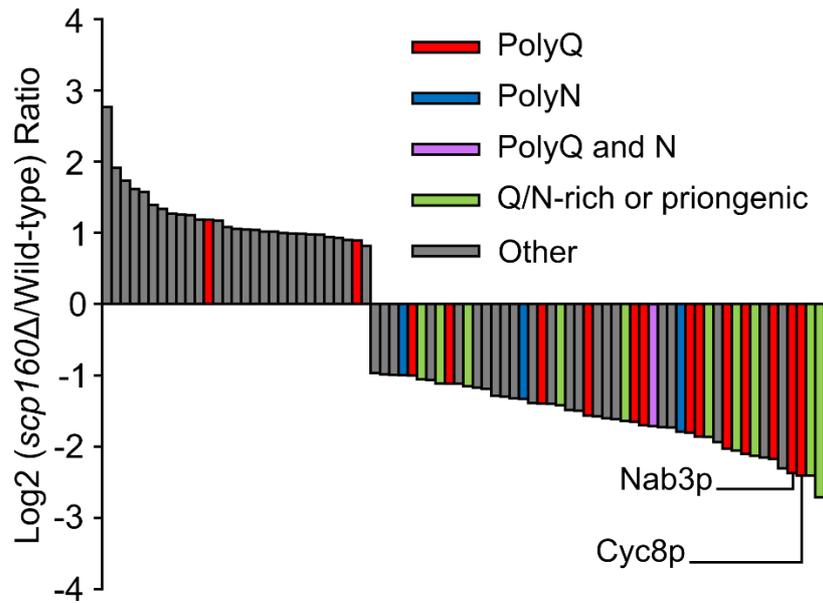
Figure 3.2.3 – Aggregation of many Q/N-rich proteins are reduced in *scp160Δ* cells. Filter-trap binding and quantitative mass spectrometry (dimethyl labeling) showed that aggregation of 77 proteins are altered in *scp160Δ* cells under normal growth conditions. Of the 48 proteins whose aggregation was decreased in *scp160Δ* cells, 27 (56.3%) contained a polyQ, polyN, or LCR as determined by manual inspection of the protein sequence or bioinformatics search from Alberti et al., 2009 and Michelitsch and Weissman, 2000.

Reduction in aggregation of two Q/N-rich proteins – Cyc8p and Nab3p – was validated using conventional filter-trap binding with immunoblotting and microscopy. Filter-trap binding showed reduced aggregation of Cyc8p and Nab3p, and western blot showed that this was not due to a difference in protein abundance between wild-type and *scp160Δ* cells (Figure 3.2.4A and B). To complement the filter-trap binding, aggregation of GFP-tagged Cyc8p and Nab3p were also assessed by microscopy. To approximate aggregation, a ratio of maximum/minimum signal intensity (signal intensity ratio) was calculated for the GFP-containing area (see Appendix, Cheng et al., 2018 for more detail). An evenly distributed signal (suggesting soluble proteins) would give a signal intensity ratio close to 1, while foci-containing signal (suggesting aggregated proteins among a soluble background) would give a higher signal intensity ratio. Our microscopic analysis showed a modest but significant downward shift of the Cyc8p-GFP ratio distribution for *scp160Δ* cells compared to wild-type cells (Figure 3.2.4C and D). The Nab3p-GFP ratio also showed a modest but insignificant downward shift for *scp160Δ* cells compared to wild-type cells.
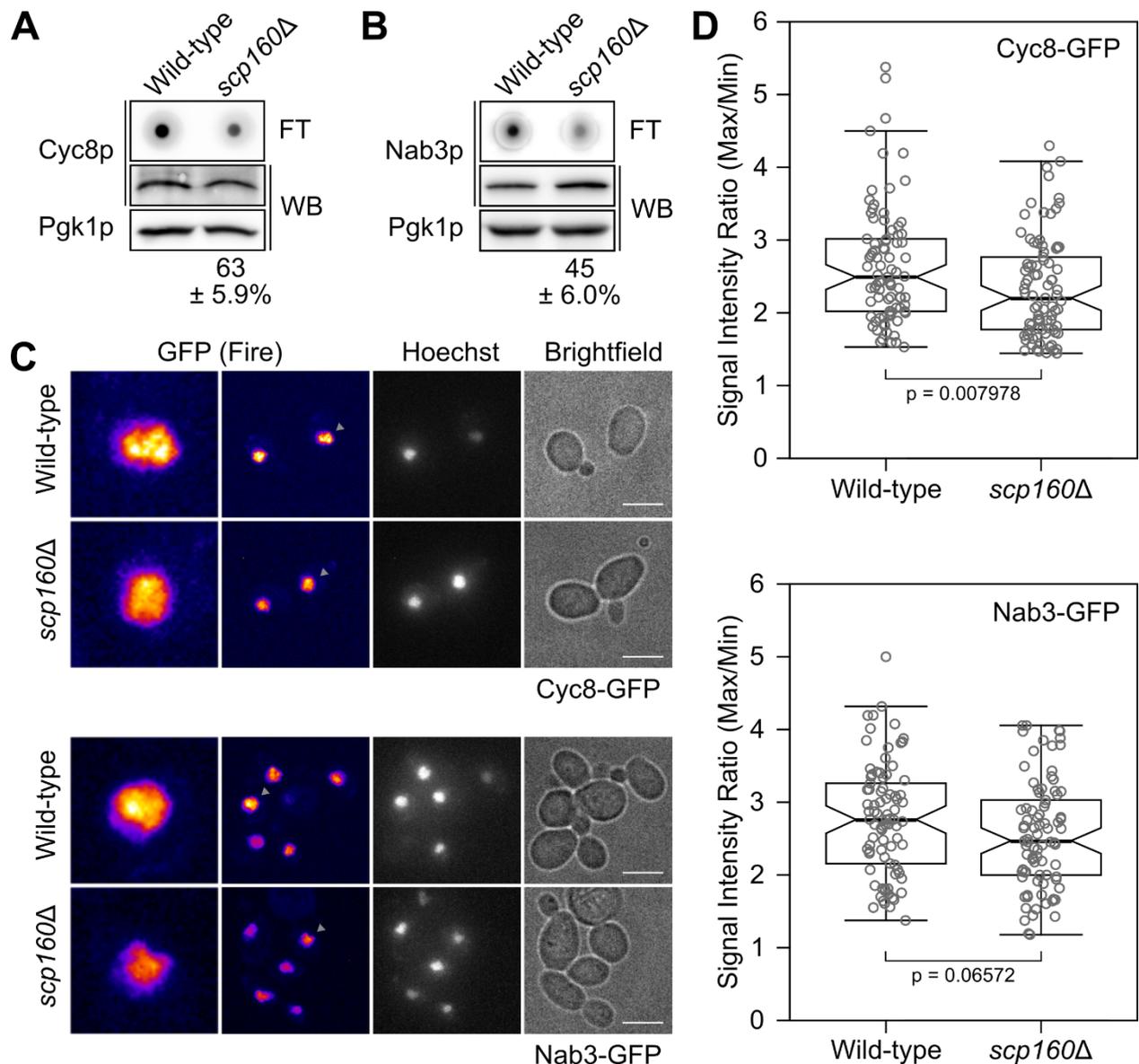
Figure 3.2.4 – Aggregation of Cyc8p and Nab3p are reduced in *scp160Δ* cells. Filter-trap binding showed that SDS-resistant aggregation of **A)** Cyc8p and **B)** Nab3p are reduced in *scp160Δ* cells and that this is not due to changes in protein abundance. The numbers show the average immunosignal for GFP (Cyc8p) or myc (Nab3p) of *scp160Δ* cells compared to wild-type cells. SEM is shown for three biological replicates. **B-C)** Microscopy of GFP-tagged Cyc8p and Nab3p showed reduced aggregation in *scp160Δ* as approximated by calculation of a maximum/minimum signal intensity ratio based on the brightest and dimmest pixels in the GFP-containing area. **D)** Distribution of the signal intensity ratio for Cyc8p-GFP and Nab3p-GFP in wild-type vs *scp160Δ* cells. Significance of the shift in ratios between wild-type and *scp160Δ* cells were tested by Mann-Whitney U test. n=81 for Cyc8p-GFP and n=85 for Nab3p-GFP.

This study demonstrates the widespread effect of Scp160p on protein aggregation mediated by Q/N-rich regions and/or low-complexity regions (LCRs). The proteins whose aggregation is influenced by Scp160p spans a variety of functions, and this may provide a hint to how Scp160p is implicated in many seemingly unrelated cellular processes.

3.3 – Plasmid-based *SCP160* does not rescue aggregation phenotypes

To ask whether plasmid-based expression of *SCP160* could restore 97QmCh aggregation in *scp160Δ* cells, I transformed either a pRS313-*SCP160* or the corresponding empty pRS313 plasmid into wild-type and *scp160Δ* cells that also carry the 97QmCh plasmid. The pRS313-*SCP160* plasmid includes the endogenous *SCP160* promoter and terminator sequences. 97QmCh was induced in the resulting cells for 3 and 7 hours, and 97QmCh aggregation was observed by microscopy as described in the Research Article. Although Scp160p was expressed from the pRS313-*SCP160* plasmid (Figure 3.3A), this did not increase 97QmCh aggregation in the *scp160Δ* background (Figure 3.3B). While it cannot be ruled out that the plasmid-based Scp160p is non-functional, the inability of this protein to restore 97QmCh aggregation may be due to irreversible changes in cell function incurred by the loss of *SCP160* which can impact protein biosynthesis. Similarly, plasmid-based expression of Scp160p was also unable to rescue the ploidy phenotype of *scp160Δ* cells (Wintersberger et al., 1995). Despite the inability of plasmid-based *SCP160* expression to restore 97QmCh aggregation, I observed a slight increase in aggregation in wild-type cells carrying the pRS313-*SCP160* plasmid (Figure 3.3B). This effect is more consistently observed for cells containing small aggregates after 3 hours of induction (78.2±2.1% of wild-type cells+pRS313-*SCP160* vs 63.2±1.5% of wild-type cells+pRS313). After 7 hours of induction, the average percentage of cells containing large aggregate(s) is higher when the wild-type cells carried the pRS313-*SCP160* plasmid (36.4±9.1% of wild-type cells+pRS313-*SCP160* vs 25.7±3.4% of wild-type cells+pRS313). However, this increase is largely attributed to the third of three replicates, as the percentages of cells containing large aggregate(s) is more comparable when excluding the one replicate (28.6±8.1% of wild-type cells+pRS313-*SCP160* vs 25.9±5.9% of wild-type cells+pRS313; data not shown). Although this difference between the wild-type cells carrying the pRS313-*SCP160* or empty vector is not statistically significant (regardless of the inclusion of the third replicate), the large variation introduced by the third replicate warrants additional analyses by microscopy and/or filter-trap binding.
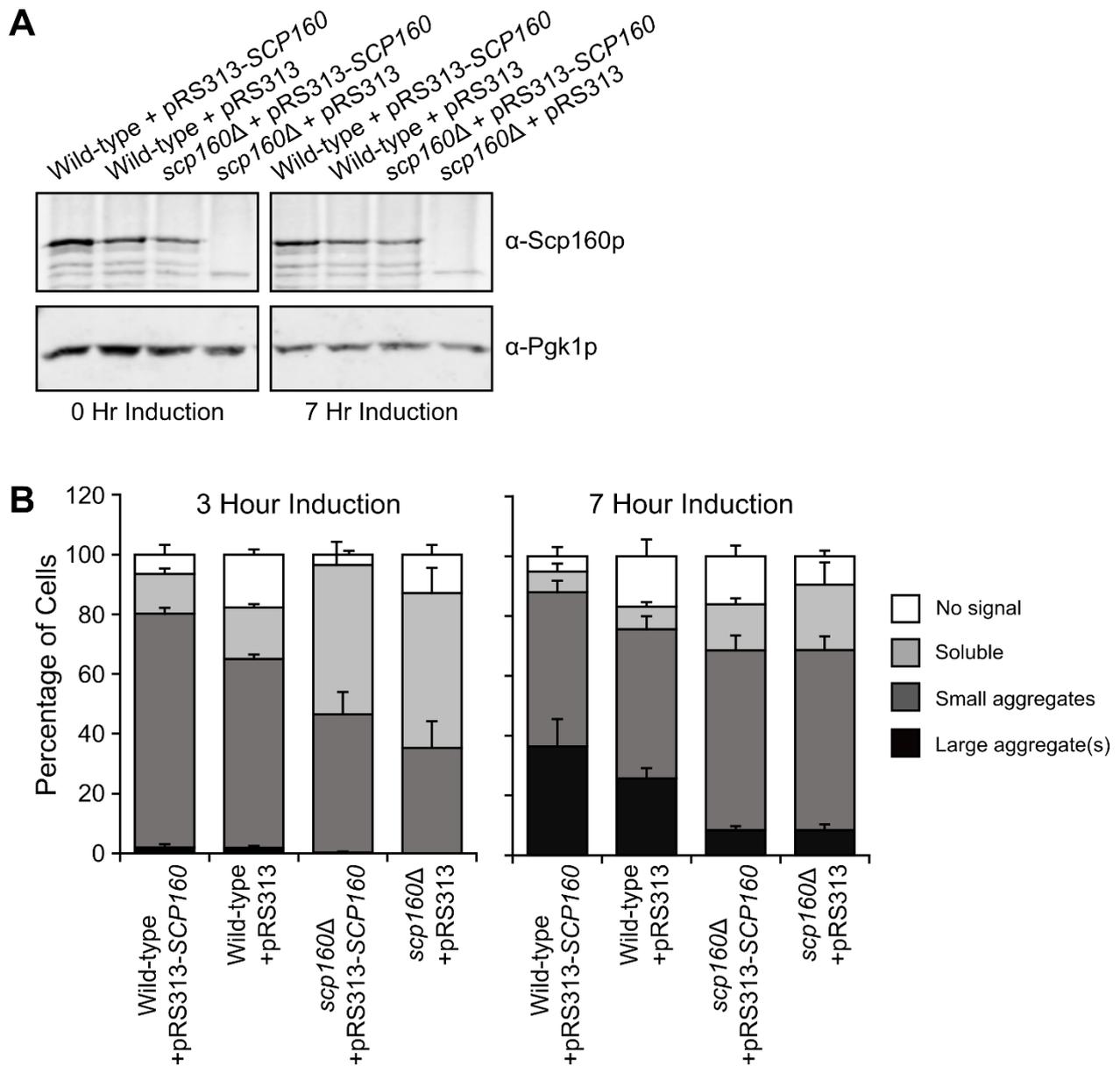
Figure 3.3 – Plasmid-based _SCP160_ do not restore 97QmCh aggregation. A) Western blot shows Scp160p expression under its endogenous promoter from a plasmid. Scp160p levels were detected at 0 and 7 hours after induction of the 97QmCh reporter with galactose. **B)** The percentages of cells containing no detectable mCherry signal (white), soluble mCherry (light grey), small aggregates (dark grey), and large aggregate(s) (black). See Research Article for more details. Plasmid-based expression of Scp160p did not restore or increase the percentage of cells containing aggregated mCherry signal after 3 or 7 hours of induction in _scp160Δ_ cells. Interestingly, plasmid-based expression of Scp160p in wild-type cells increased the percentage of cells containing aggregated mCherry signal after 3 and 7 hours of induction. Error bars represent SEM of three biological replicates.

3.4 – Scp160p does not interact with the 97QmCh reporter protein

To ask whether Scp160p might interact with and therefore influence aggregation of the 97QmCh reporter protein, I assessed if 97QmCh coimmunoprecipitates with Scp160p-GFP. To this end, I immunoprecipitated Scp160p-GFP from lysates in which 97QmCh expression was induced for 7 hours.

Figure 3.4A shows that although Scp160p-GFP was pulled down, the 97QmCh reporter protein did not coimmunoprecipitate. Conversely, pulldown of 97QmCh also failed to coimmunoprecipitate Scp160p (Figure 3.4B). Together, these results suggest that Scp160p does not stably interact with 97QmCh and that Scp160p's influence on the aggregation of the polyQ reporter proteins is unlikely to come from direct protein-protein interaction.
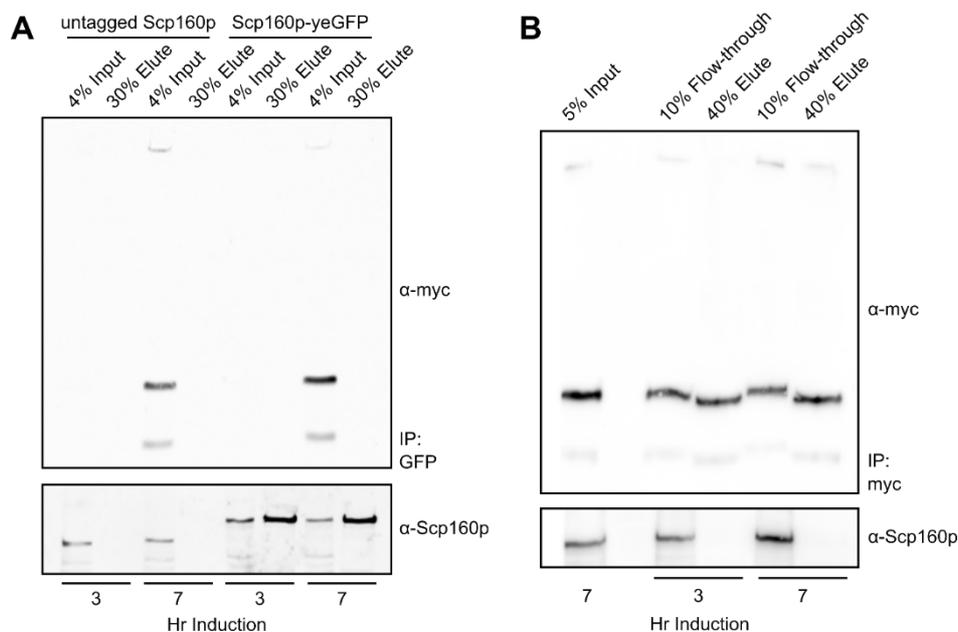
Figure 3.4 – Scp160p and 97QmCh does not interact physically. A) Pulldown of GFP-tagged Scp160p was unable to coimmunoprecipitate the 97QmCh reporter protein (probed against its N-terminal myc tag). **B)** Pulldown of 97QmCh via its N-terminal myc tag was unable to coimmunoprecipitate endogenous Scp160p.

3.5 – KH13-14 domains contribute to polyQ aggregation

Domains KH13 and 14 of Scp160p are important for the polysome-associated and mRNA-binding functions of Scp160p (Hirschmann et al., 2014; Li et al., 2004), but not for ploidy maintenance (see 3.1.4). Therefore, I wanted to test if these domains could be involved in facilitating aggregation of the 97QmCh reporter. Using the C-terminal truncation Scp160p(KH1-12)-GFP described above, I assessed 97QmCh aggregation by filter-trap binding. Strikingly, loss of KH13-14 (Scp160p(KH1-12)-GFP) reduced aggregation of 97QmCh compared to the full-length protein (Figure 3.5A). Western blotting showed that the abundance of 97QmCh reporter protein was comparable between Scp160p(KH1-12)-GFP, wild-type, and *scp160Δ* cells (Figure 3.5B). Although these observations were made with only two biological replicates, they suggest that KH13-14 is involved in the aggregation of 97QmCh. Moreover, given the importance of KH13-14 in Scp160p's association with polysomes, this may link

Scp160p's role in translation with polyQ-mediated protein aggregation. However, it is likely that the other KH domains 1-12 also contribute to Scp160p's role in polyQ-mediated aggregation as the effect of the Scp160p(KH1-12) truncation is less severe than *scp160Δ.* Finally, since the loss of KH13-14 did not affect ploidy maintenance, this observation along with that using the Tet-off *SCP160* system (Cheng et al., 2018), argue that the aggregation phenotype of *scp160Δ* cells is independent of the ploidy phenotype.
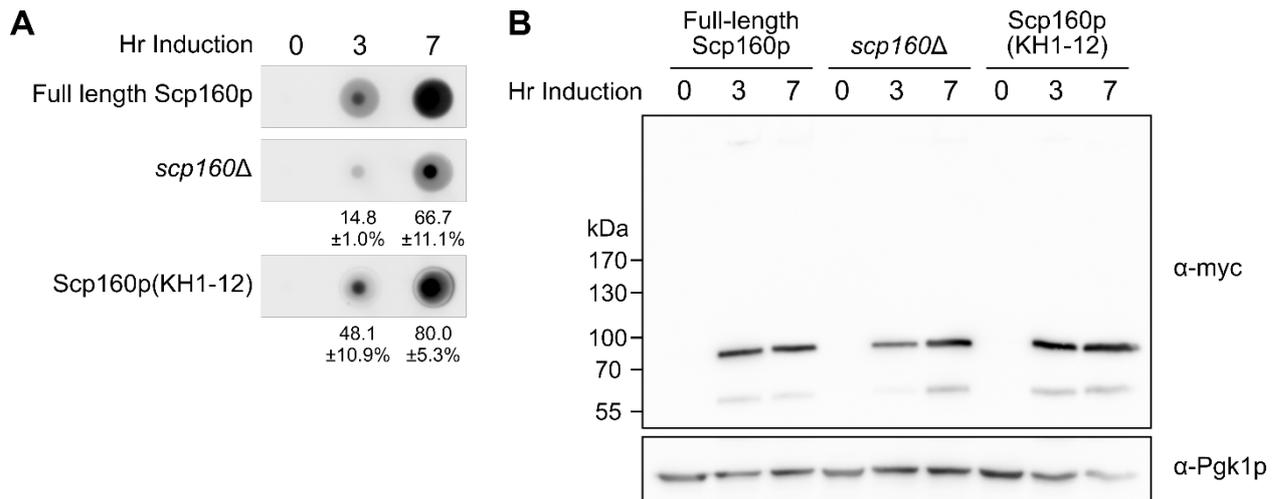
Figure 3.5 – KH13-14 contributes to 97QmCh aggregation. A) Filter-trap binding show that the aggregation of 97QmCh is reduced in cells expressing a C-terminally truncated Scp160p(KH1-12)-GFP. The numbers show the average immunosignal for myc (97QmCh) of either *scp160Δ* or Scp160p(KH1-12)-GFP cells compared to wild-type cells. SEM is shown for two biological replicates. **B)** Western blot shows similar abundance of 97QmCh protein between wild-type, *scp160Δ*, and Scp160p(KH1-12)-GFP cells after 7 hours of induction.

3.6 – Reduced polyQ aggregation appears independent of protein quality control

3.6.1 – Cellular levels of many chaperones remain comparable between wild-type and scp160Δ cells

Molecular chaperones have been reported to play roles in disaggregation of polyQ aggregates (Sakahira et al., 2002). Therefore, I checked whether the steady-state levels of several chaperones differed between wild-type and *scp160Δ* cells. The protein levels of select chaperones belonging to various families were assayed by western blot. These chaperones included the small heat shock proteins (sHSP) Hsp26p and Hsp42p, the HSP40 family members Sis1p and Ydj1p, the HSP70 member Ssa1p, the HSP90 members Hsp82p and Sti1p, and Hsp104p. The levels of these chaperones were comparable between wild-type and *scp160Δ* cells (Figure

3.6.1A). Thus, the reduction in Q/N-rich-mediated protein aggregation cannot be explained by increased steady-state levels of the chaperones tested.
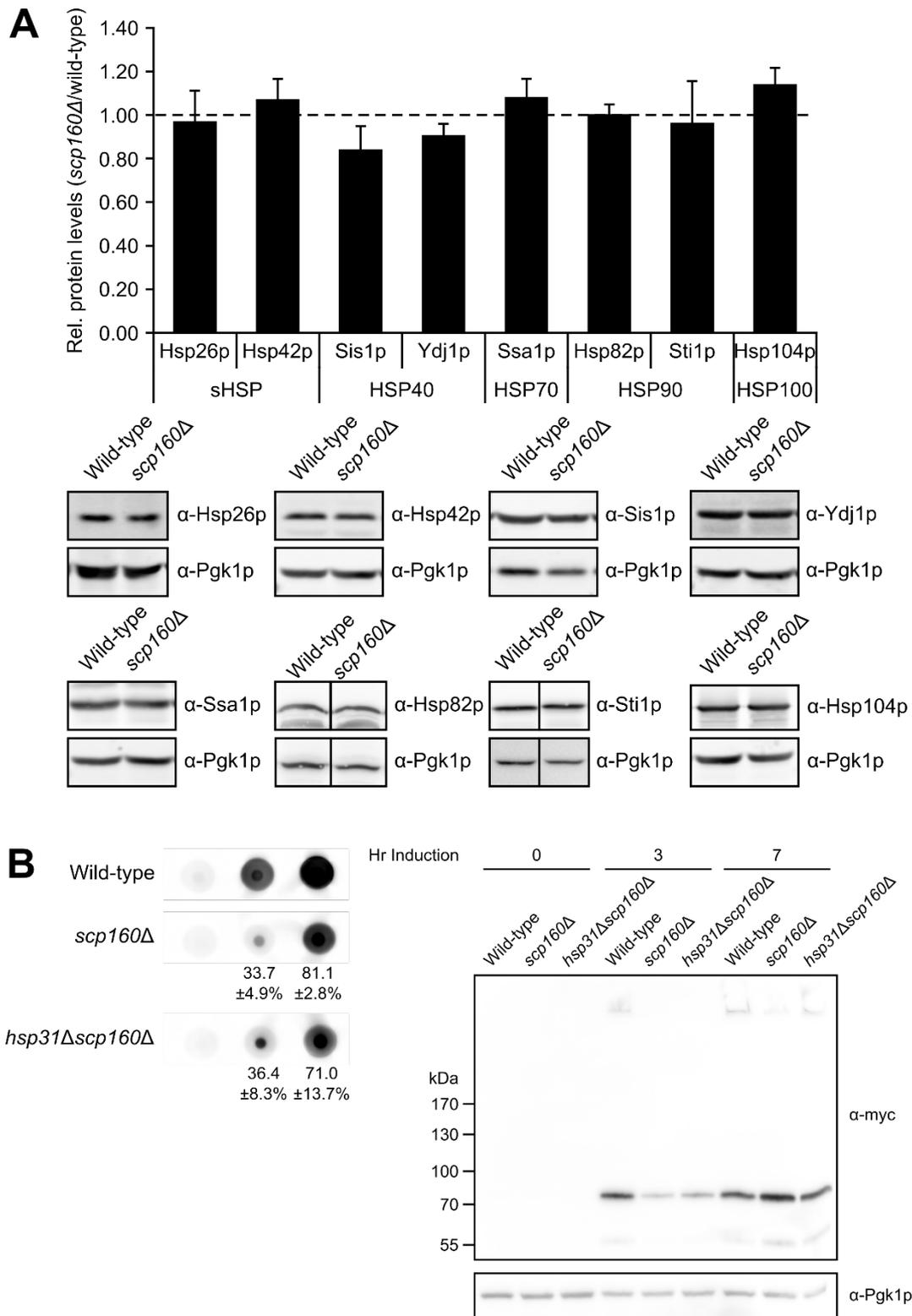
Figure 3.6.1 – Similar abundance of chaperones in wild-type and *scp160Δ* cells. A) Protein abundance of several chaperone family members are comparable between wild-type and *scp160Δ* cells. **B)** 97QmCh aggregation is not restored in *hsp31Δscp160Δ* cells, suggesting that reduced aggregation in *scp160Δ* is unlikely to function via Hsp31p.

While analyzing data from the combined filter-trap and quantitative mass spectrometry, I noticed that the protein level of Hsp31p in the whole cell lysate was approximately 1.78 times more abundant in *scp160Δ* cells compared to wild-type cells (1.73 times in replicate 1 and 1.78 times in replicate 2). Hsp31p is a stress-inducible chaperone and member of the conserved DJ-1 superfamily (Aslam and Hazbun, 2016). Notably, Hsp31p has been reported to function with Hsp104p in *S. cerevisiae* to prevent formation of the [PSI+] prion (Aslam et al., 2016) and disaggregate α-synuclein fibrils *in vitro* (Tsai et al., 2014). In order investigate if the change in Hsp31p levels might account for the reduced aggregation of 97QmCh in *scp160Δ* cells, I generated an *hsp31Δscp160Δ* double mutant. 97QmCh expression was induced in these cells for 3 and 7 hours, and aggregation was assessed by filter-trap binding. This analysis showed that the level of 97QmCh aggregation was comparable between *scp160Δ* and *hsp31Δscp160Δ* cells (Figure 3.6.1B). The amount of 97QmCh trapped on the membrane from *scp160Δ* lysates was 33.7±4.9% that of wild-type after 3 hours of induction and 81.1±2.8% after 7 hours of induction. Similarly, the amount of 97QmCh trapped from *hsp31Δscp160Δ* lysate was 36.4±8.3% that of wild-type after 3 hours of induction and 71.0±13.7% after 7 hours of induction. Although it cannot be ruled out that loss of both *SCP160* and *HSP31* leads to changes in cellular homeostasis that include upregulation of other chaperones and protein quality control systems, this observation suggests that reduced polyQ aggregation in *scp160Δ* cells does not function through Hsp31p.

3.6.2 – 97Q(CAG)mCh unlikely to be subject to ribosome-associated quality control

In addition to the cellular chaperones such as the heat shock proteins, the ribosome is also a site for monitoring mRNA and protein quality known as ribosome-associated quality control (RQC) (Brandman and Hegde, 2016). Stalling of the ribosome due to mRNA secondary structures or truncation, as well as limited cognate tRNAs can trigger ribosome-associated quality control pathways like no-go decay (NGD) and non-stop decay (NSD) (Brandman and Hegde, 2016; Matsuda et al., 2014; Shoemaker and Green, 2012). These pathways release the stalled ribosomes and target the aberrant mRNA and its nascent peptidyl-tRNA for degradation, thereby protecting the cell from potentially damaging protein products (Matsuda et al., 2014). In *S. cerevisiae* the Dom34p-Hbs1p complex has been reported to trigger NGD and

NSD by splitting the stalled ribosomes, and Dom34p has been proposed as the endonuclease which cleaves the mRNA at the stall site (Brandman and Hegde, 2016; Doma and Parker, 2006).

The polyQ-encoding region of the 97Q(CAG)mCh reporter is a long CAG trinucleotide repeat, which is able to form a long stem-loop by self-complementation (Nalavade et al., 2013). Moreover in *S. cerevisiae*, CAG is the infrequent glutamine-encoding codon and its cognate tRNA is expressed in low-copy number. These two attributes related to the 97Q(CAG)mCh reporter makes it a potential target for RQC. Indeed, during my analyses, I noticed that the protein abundance of 97Q(CAG)mCh was much lower in wild-type cells compared to *scp160Δ* cells (Figure 3.6.2A and Research Article), which could reflect its targeting by RQC. To investigate whether the 97Q(CAG)mCh mRNA may be subject to NGD, I assessed the steady-state levels 97Q(CAG)mCh mRNA at 3, 5, and 7 hours after induction in wild-type, *scp160Δ* and *dom34Δ* cells (Figure 3.6.2B). If 97Q(CAG)mCh is subject to NGD in wild-type cells, I should observe an increase in steady-state mRNA levels in *dom34Δ* cells.

In *scp160Δ* cells, 97Q(CAG)mCh mRNA was induced to much greater (but not statistically significant, due to high variance between samples) levels after just 3 hours of induction compared to wild-type cells. After 5 and 7 hours of induction the level of 97Q(CAG)mCh was significantly higher in *scp160Δ* cells than wild-type cells (7.0±1.5 fold and 7.4±0.9 fold over wild-type). Interestingly, the induction of 97Q(CAG)mCh mRNA in *dom34Δ* cells had a different kinetic than wild-type and *scp160Δ* cells. 97Q(CAG)mCh mRNA level in *dom34Δ* cells displayed an initial (not statistically significant) increase over wild-type cells, like that of *scp160Δ* cells, after 3 hours of induction. However, the levels of steady-state level mRNA after 5 and 7 hours of induction in *dom34Δ* cells were only 1.5±0.2 fold and 1.6±0.2 fold over that in wild-type cells. These observations suggest that the difference in 97Q(CAG)mCh mRNA levels between wild-type and *scp160Δ* cells cannot be solely explained by NGD. In support, western blot showed that the 97Q(CAG)mCh reporter protein is much less abundant in both wild-type and *dom34Δ* cells as compared to *scp160Δ* cells (Figure 3.6.2A).
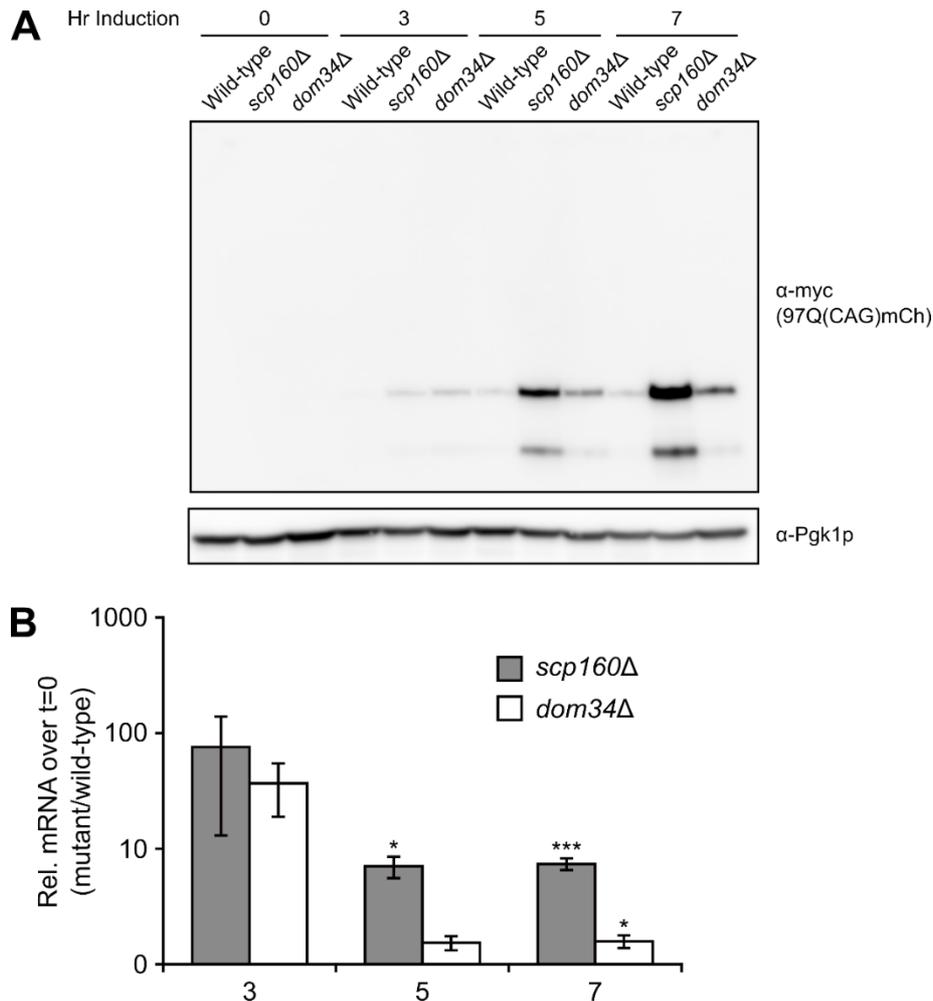
Figure 3.6.2 – 97Q(CAG)mCh mRNA is more abundant in *scp160Δ* cells. A) Western blot shows that 97Q(CAG)mCh protein levels are more abundant in *scp160Δ* cells. 97Q(CAG)mCh protein levels are modestly higher in *dom34Δ* cells compared to wild-type cells, but do not reach levels observed in *scp160Δ* cells. **B)** RT-qPCR show that 97Q(CAG)mCh mRNA is more abundant in *scp160Δ* cells compared to wild-type cells. 97Q(CAG)mCh mRNA levels are also modestly increased over wild-type in *dom34Δ* cells. Error bars represent SEM of three biological replicates. * $p < 0.05$, *** $p < 0.005$.

3.6.3 – Loss of the RQC component Ltn1p has little effect on 97QmCh aggregation

After the splitting of stalled ribosomes by Dom34p-Hbs1 in RQC, the aberrant nascent peptidyl-tRNA-60S complex is targeted for proteasomal degradation by the concerted action of Rqc2p, Ltn1p, Rqc1p, and Cdc48p (Brandman and Hegde, 2016; Brandman et al., 2012; Ikeuchi and Inada, 2016). Together, nascent peptides from stalled ribosomes are marked by non-templated addition of C-terminal adenines and threonines (CAT-tails) and poly-ubiquitination (Brandman and Hegde, 2016; Shen et al., 2015). Poly-ubiquitination is mediated by Ltn1p at the exit-tunnel of the 60S ribosomal subunit (Brandman and Hegde, 2016). Recently, it was reported that collision of ribosomes on the mRNA caused by stalling can trigger NGD (Simms et al.,

2017). Our lab has proposed a model for Scp160p in which it facilitates recycling of tRNAs by spanning subsequent ribosomes on an mRNA (see section 1.4). In such a scenario, loss of Scp160p might also increase ribosome collision and trigger NGD. As such, I wanted to explore this possibility and ask whether the reduced aggregation of my 97QmCh reporter protein in *scp160Δ* cells might reflect greater Ltn1p-mediated poly-ubiquitination and subsequent clearance of 97QmCh. Indeed, loss of Ltn1p has been reported to disrupt RQC resulting in increased aggregation of polybasic nascent-chains and of non-stop mRNAs (Choe et al., 2016). To this end, I assessed 97QmCh protein levels and aggregation in *ltn1Δ* and *scp160Δltn1Δ* cells.
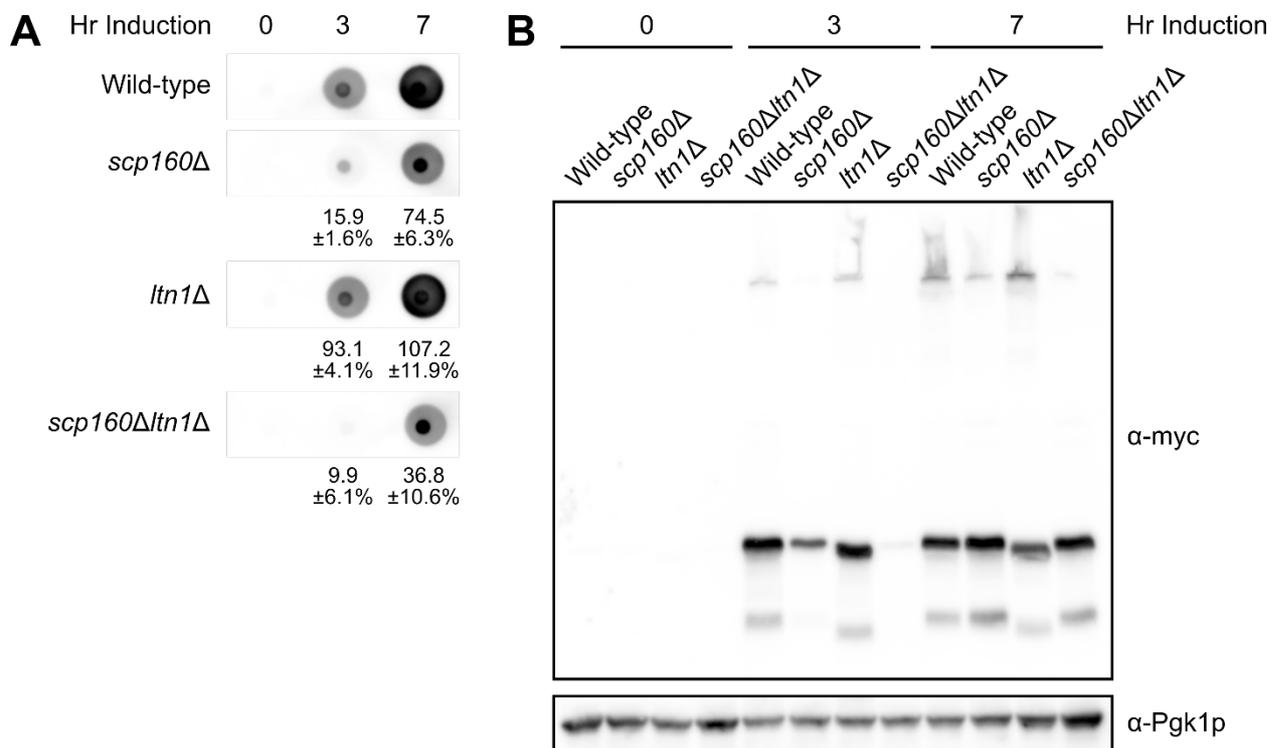
Figure 3.6.3 – 97QmCh is further reduced in an *scp160Δltn1Δ* strain. A) Filter-trap binding shows that the level of 97QmCh aggregation is negligibly affected in *ltn1Δ* cells but is further reduced in *scp160Δltn1Δ* cells. The numbers show the average immunosignal for myc (97QmCh) of either *scp160Δ*, *ltn1Δ*, or *scp160Δltn1Δ* cells compared to wild-type cells. SEM is shown for three biological replicates. **B)** Western blot shows similar abundance of 97QmCh protein between wild-type, *scp160Δ*, and *ltn1Δ* cells after 7 hours of induction. The accumulation of 97QmCh is slower in *scp160Δltn1Δ* cells.

Consistent with previous observations by filter-trap binding, 97QmCh aggregation was reduced in *scp160Δ* cells (15.9±1.6% and 74.5±6.3% of wild-type after 3 and 7 hours induction, respectively). In *ltn1Δ* cells, 97QmCh was aggregated to a similar degree as in wild-type cells (93.1±4.1% and 107.2±11.9% of wild-type after 3 and 7 hours induction, respectively) (Figure 3.6.3A). It should be noted that this quantification may underestimate the level of aggregation of 97QmCh in *ltn1Δ* cells as western blot shows that the abundance of the monomeric reporter protein (in the

separating gel) is lower than in wild-type cells while the amount of aggregated material that remains trapped in the loading pocket is higher (Figure 3.6.3B). This is in agreement with the role of Ltn1p in polyubiquitinating aberrant proteins in RQC. Strikingly, 97QmCh aggregation is drastically reduced in the *scp160Δltn1Δ* cells (9.9±6.1% and 36.8±10.6% of wild-type after 3 and 7 hours of induction, respectively) (Figure 3.6.3A). Western blot shows that the relative abundance of the 97QmCh protein is lower in the *scp160Δltn1Δ* cells compared to the other cell types. This is especially noticeable after 3 hours of induction (Figure 3.6.3B). Together, these observations suggest that despite the interacting effects of deleting *SCP160* and *LTN1*, the reduction of 97QmCh aggregation upon loss of Scp160p functions separately from Ltn1p.

3.7 – Reduced polyQ aggregation may be associated with aberrant translation elongation speed

In vitro and *in vivo* studies have demonstrated that codon usage and, by extension, translation elongation kinetics can impact co-translational folding and ultimately function of the nascent peptide (Buhr et al., 2016; Fu et al., 2016; Pechmann et al., 2014). Although the data presented above suggests that polyQ-mediated protein aggregation is reduced in *scp160Δ* cells regardless of codon usage, it remains a possibility that loss of Scp160p perturbs elongation kinetics of Q/N-rich regions that could alter aggregation. Indeed, the proposed role for Scp160p in facilitating the reuse of tRNAs cognate to anticorrelated codons suggests that elongation speed could be negatively influenced in Scp160p's absence. To ask if polyQ-mediated aggregation can be reduced by slowing translation elongation, I treated wild-type and *scp160Δ* cells with a mild dose (30ng/mL) of cycloheximide and assessed 97QmCh aggregation by microscopy and filter-trap binding. Cycloheximide is a commonly used antibiotic that inhibits translation elongation by blocking the translocation step (McKeehan and Hardesty, 1969; Schneider-Poetsch et al., 2010). Importantly, mild cycloheximide treatment has been shown to be able to slow (but not completely inhibit) protein synthesis and promote proper protein folding (Meriin et al., 2012; Simms et al., 2017). During the culturing and induction period, I noticed that although this dosage of cycloheximide did slow growth modestly, the populations remained viable.
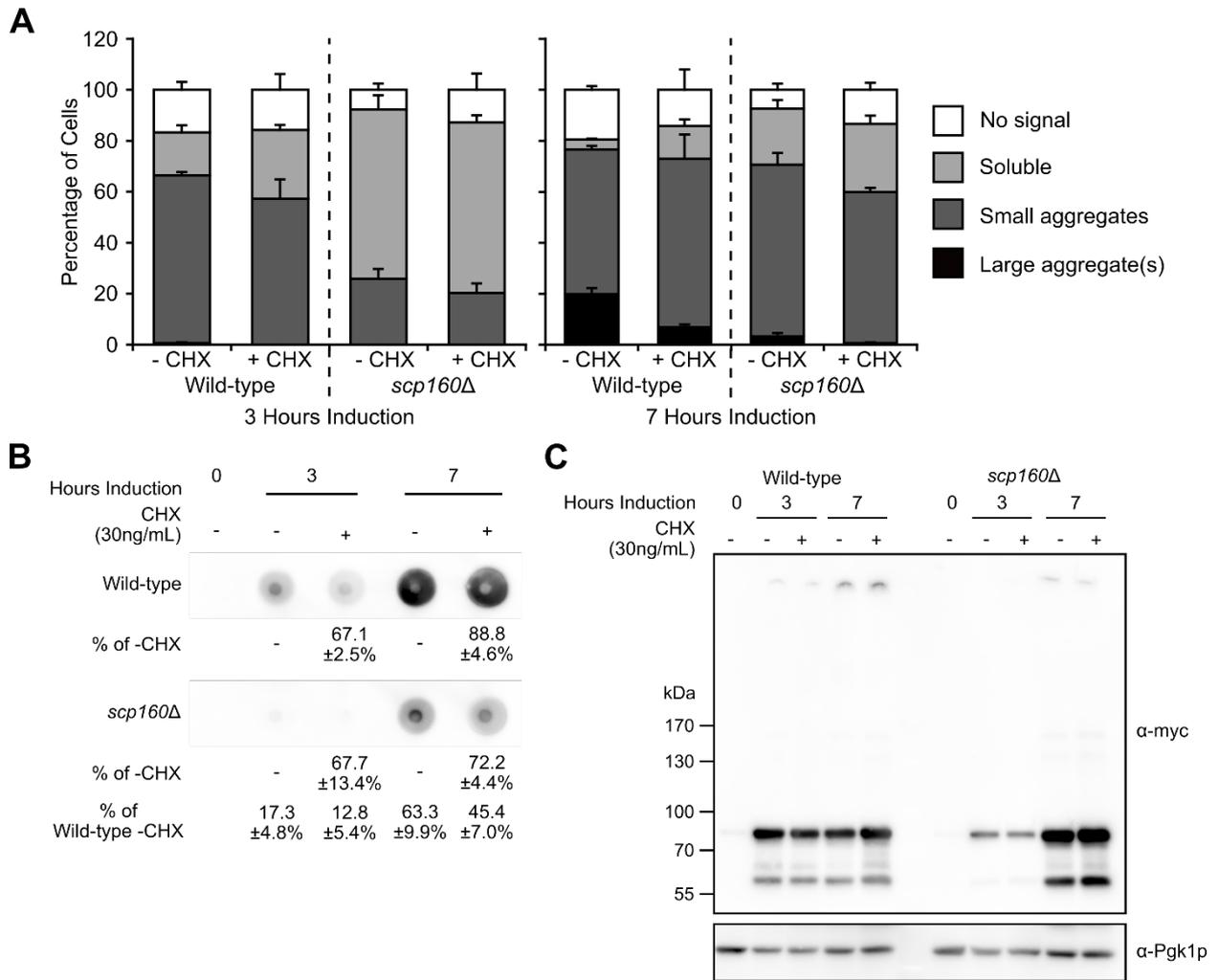
Figure 3.7 – Mild cycloheximide treatment reduces 97QmCh aggregation. A) The percentages of cells containing no detectable mCherry signal (white), soluble mCherry (light grey), small aggregates (dark grey), and large aggregate(s) (black). See Research Article for more details. Error bars represent SEM of three biological replicates. **B)** Filter-trap binding shows that perturbation of translation elongation by treatment with 30ng/mL cycloheximide lead to the reduction of a 97QmCh. Numbers listed in the row "% of –CHX" show the average immunosignal of the cycloheximide-treated cells as a percentage of the corresponding untreated cells. Numbers listed in the row "% of Wild-type –CHX" show the average immunosignal of the *scp160Δ* cells (treated and untreated) as a percentage of the corresponding untreated wild-type cells. SEM of three biological replicates are shown. **C)** Western blot shows that the overall protein abundance of 97QmCh is not affected by the mild cycloheximide treatment.

Microscopy showed that mild cycloheximide treatment modestly reduce the percentage of cells with small aggregates after 3 hours of induction (2 hours cycloheximide treatment) (Figure 3.7A). The effect was stronger after 7 hours of induction (6 hours of cycloheximide treatment) with a reduction of treated cells containing large aggregate(s) (in wild-type cells, 19.8±2.3% untreated vs 6.9±1.05% treated). Notably, the percentage of treated wild-type cells that contain large aggregate(s) at 7 hours of induction approached the percent of untreated *scp160Δ* cells (6.9±1.5% vs 3.3±1.3%). The treatment of *scp160Δ* cells with cycloheximide also

led to a reduction in the percentage of cells with large aggregate(s) (3.3±1.3% untreated vs 0.6±0.3% treated) and small aggregates (67.3±4.6% untreated vs 59.4±1.7% treated) after 7 hours of induction.

This data is consistent with previous studies in which perturbation of elongation rate can reduce stress-induced polyQ aggregation in *C. elegans* (Kim and Strange, 2013) and formation of aggresomes in HeLa cells . However, attempts to assess Cyc8p-GFP aggregation under mild cycloheximide treatment has yielded variable effects, making it difficult to draw a conclusion. The variability in this regard may lie in the relatively low abundance of endogenous Cyc8p and/or reflect a need to further optimize cycloheximide treatment conditions.

3.8 – Loss of Bfr1p influences polyQ reporters differently than loss of Scp160p

3.8.1 – All three polyQ reporters are toxic to bfr1Δ cells

Due to the close functional relation between Scp160p and Bfr1p, I investigated in parallel how loss of Bfr1p might influence the three polyQ reporters described above (Figure 2). In contrast to *scp160Δ* cells, all three polyQ reporters conferred severe toxicity to *bfr1Δ* cells, with the 97QmCh and 104Q(CAA)mCh reporters similarly toxic to *bfr1Δ* mutants as the 97Q(CAG)mCh reporter (Figure 3.8.1). This difference in toxicity of the reporters between *scp160Δ* and *bfr1Δ* mutants suggests Scp160p and Bfr1p likely play different roles in polyQ-induced toxicity
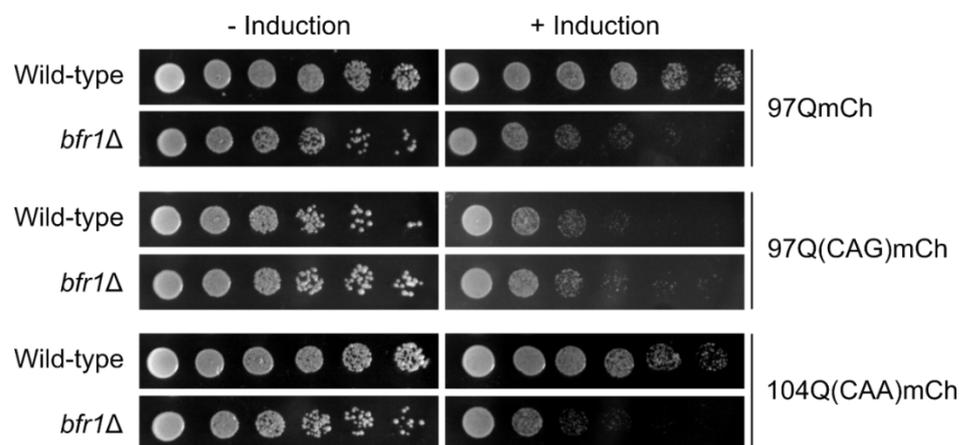
Figure 3.8.1 – All three polyQ reporter are toxic to *bfr1Δ* cells. Serial dilution assays show that the induction of the polyQ reporter proteins confer severe toxicity to *bfr1Δ* cells regardless of codon usage. This contrasts with *scp160Δ* cells in which only the 97Q(CAG)mCh reporter conferred severe toxicity.

3.8.2 – Induction of 97QmCh reporter mRNA is poor in bfr1Δ mutants

To assess how polyQ-mediated aggregation might be affected in *bfr1Δ* cells, I followed 97QmCh aggregation by microscopy and filter-trap binding. I chose to

observe 97QmCh because it is well expressed and a good candidate for preliminary analysis given its mixed codon usage, but also because its expression conferred a high level of toxicity in *bfr1Δ* cells that was not observed in wild-type and *scp160Δ* cells. Microscopic analysis was performed as described in Cheng et al., 2018, in which cells were categorized into 4 classes based on the distribution of the mCherry signal. Strikingly, even after 15 hours induction about half the amount of *bfr1Δ* cells displayed no detectable mCherry signal (62.7±5.8% after 3 hours induction, 66.9±3.6% after 7 hours induction, and 54.9±6.6% after 15 hours induction) (Figure 3.8.2A). This is in stark contrast to wild-type and *scp160Δ* cells, in which only a small percentage of cells showed no detectable mCherry signal (9.3±4.7% after 3 hours induction, 14.8%±4.7% after 7 hours induction, and 12.1±1.2% after 15 hours induction in wild-type cells; 20.3±2.9% after 3 hours induction, 5.3±1.4% after 7 hours induction, and 3.3±1.8% after 15 hours induction in *scp160Δ* cells). This made it difficult to analyze 97QmCh aggregation by microscopy.

As a complementary assay, 97QmCh aggregation was assayed by filter-trap binding with lysates from wild-type, *scp160Δ*, and *bfr1Δ* cells in which reporter expression was induced for 3, 7, and 15 hours. Quantification of the resulting immunosignals showed the amount of trapped 97QmCh from *bfr1Δ* lysates is lower than that from wild-type and *scp160Δ* lysates at each corresponding time point (Figure 3.8.2B). However, western blot of the lysates showed that in the case of *bfr1Δ* mutants, the reduced 97QmCh immunosignal is due to reduced expression of the reporter protein and not reduced aggregation (Figure 3.8.2C).

Since Bfr1p is a part of a polysome-associated mRNP complex (Lang et al., 2001), the low level of 97QmCh reporter protein in *bfr1Δ* cells may be caused by aberrant translation, mRNA stability, or transcription. To investigate these possibilities, the steady-state mRNA of 97QmCh was assessed by RT-qPCR in wild-type, *scp160Δ*, and *bfr1Δ* cells after 0, 3, 7, and 15 hours of induction. 97QmCh mRNA levels were checked using primers that targeted the mCherry-encoding nucleotide sequence. mRNA abundance relative to the level before induction was calculated using the $\Delta\Delta C_t$ method using 18S (Figure 3.8.2D) and *ACT1* (data not shown) as loading controls. RT-qPCR analyses showed that the steady-state mRNA level of 97QmCh in *bfr1Δ* cells was significantly lower than wild-type and *scp160Δ* cells at the corresponding time points (unpaired two-tailed t-test, $p<0.01$). While it cannot be ruled out that the 97QmCh reporter is more efficiently targeted for degradation in *bfr1Δ* cells, the RT-

qPCR analysis suggests that the low abundance of 97QmCh protein in these cells is not primarily due to aberrant translation.
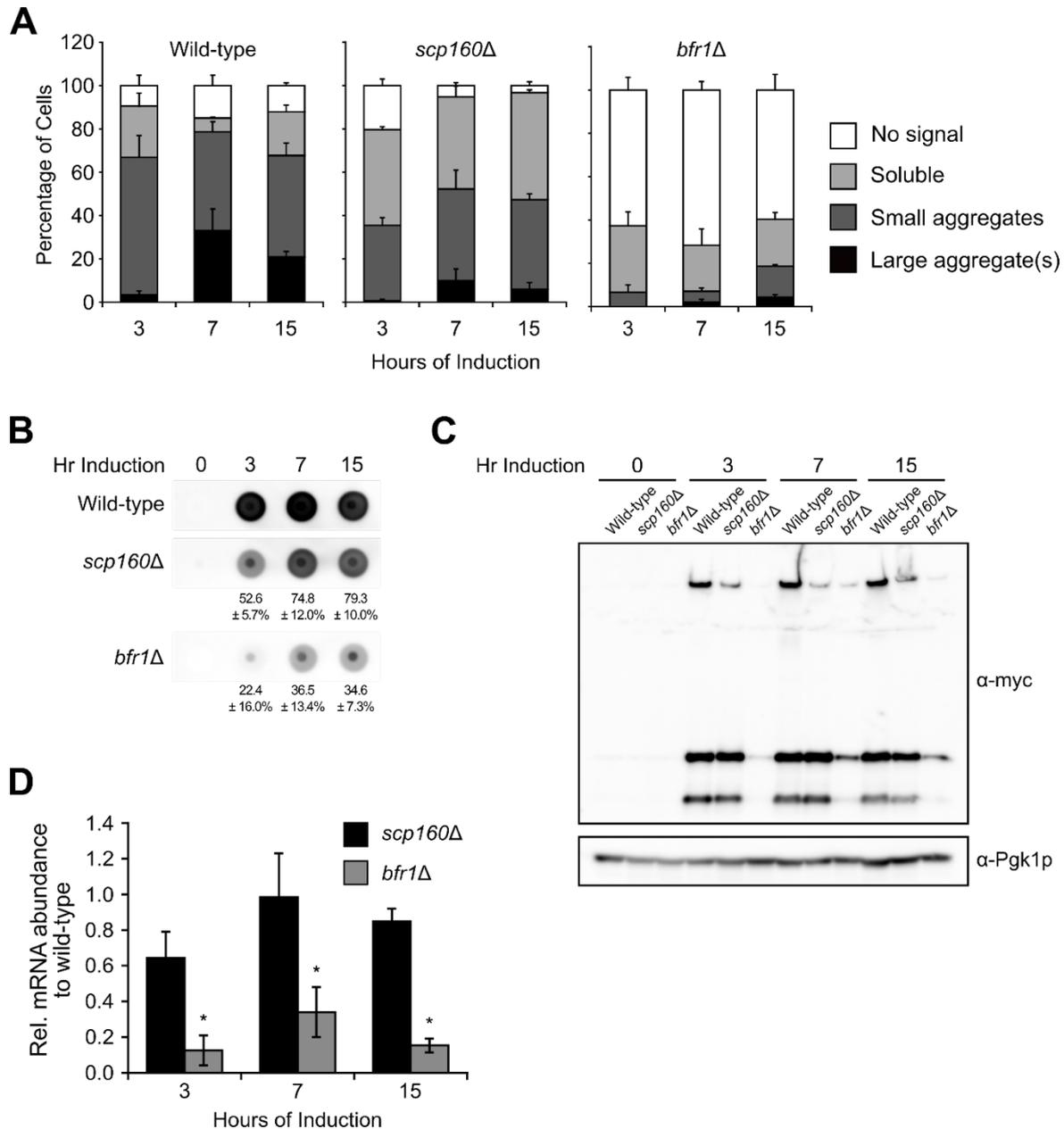
Figure 3.8.2 – Induction of polyQ reporters is poor in *bfr1Δ* cells. A) The percentages of cells containing no detectable mCherry signal (white), soluble mCherry (light grey), small aggregates (dark grey), and large aggregate(s) (black). See Research Article for more details. Unlike wild-type and *scp160Δ* cells, no signal was detected in the majority of *bfr1Δ* cells. Error bars represent SEM of three biological replicates. **B)** Filter-trap binding revealed that 97QmCh aggregation appears to be lower in *bfr1Δ* cells compared to wild-type and *scp160Δ* cells. The numbers show the average and SEM of three biological replicates. **C)** Western blot shows that the abundance of the 97QmCh reporter protein in *bfr1Δ* is much lower than that in wild-type and *scp160Δ* cells. **D)** RT-qPCR against the reporter's mCherry-encoding sequence show a significantly lower steady-state amount of 97QmCh mRNA in all three induction time points for *bfr1Δ* cells compared to wild-type cells. Error bars represent SEM of three biological replicates and * indicates p<0.01. Steady-state levels of 97QmCh mRNA at all three time points for *scp160Δ* cells were not significantly different from that of wild-type cells.

4. DISCUSSION

4.1 – Functional relationship between Scp160p and Bfr1p

4.1.1 – Scp160p and Bfr1p in ploidy maintenance

The RBPs Scp160p and Bfr1p in *S. cerevisiae* share a close functional relationship (Hogan et al., 2008; Jackson and Képès, 1994; Lang et al., 2001; Sezen et al., 2009; Weidner et al., 2014; Wintersberger et al., 1995). One of the more striking phenotype associated with *scp160Δ* and *bfr1Δ* cells is an increase in ploidy (Jackson and Képès, 1994; Wintersberger et al., 1995). In this study, I explored the relationship of Scp160p and Bfr1p in ploidy maintenance as well as protein biosynthesis of the polyQ reporters. The loss of both *SCP160* and *BFR1* did not lead to a more severe ploidy phenotype than the single deletion strains, suggesting that the two RBPs function together in ploidy maintenance. This is consistent with the roles of Scp160p and Bfr1p as translation regulators in the SESA network, which has been linked to SPB duplication (Sezen et al., 2009). However, it remains unclear how Scp160p and Bfr1p function in the SESA network in sensing SPB duplication defects and the selective translation repression of *POM34*. Furthermore, given the involvement of the *Drosophila* vigilin in heterochromatin formation and DNA fidelity (Huertas et al., 2004; Wang et al., 2005; Zhou et al., 2008), Scp160p and Bfr1p may play additional roles unrelated to the SESA network in ploidy maintenance.

To investigate if Scp160p and Bfr1p might act on other steps in ploidy maintenance, it would be of interest to monitor DNA segregation and the onset of ploidy increase upon depletion these RBPs. Such analyses would be possible by combining fluorescent protein-based chromosome labelling (described in Kitamura et al., 2007 and in a Master's thesis by Dennis Clement) and an auxin-based inducible protein depletion system (described in Nishimura and Kanemaki, 2014 and established in a Master's project module by Lisa Heinold). A comparison of the onset of ploidy (measured for example by number of divisions after protein depletion) increase upon depletion of Scp160p, Bfr1p, or both RBPs would provide insight into the hierarchical function of these proteins. Moreover, by tracking the movement of the labelled chromosomes, it may be possible to determine if the ploidy phenotype of *scp160Δ* and *bfr1Δ* cells occurs due to unchecked DNA replication and/or defective nuclear segregation. Notably, the ploidy analysis described in this work shows that KH11-14 – shown to be important for polysome-association and binding of many

mRNAs (Hirschmann et al., 2014; Li et al., 2004) – are not required to maintain exact ploidy. Identification of proteins that interact with Scp160p's KH1-10 by co-immunoprecipitation or proximity-labelling will provide clues to Scp160p's function in ploidy maintenance. The full-length Scp160p control in such an approach could also help map certain roles of the protein to either KH1-10 or KH11-14. Interestingly, none of the mutants observed by flow cytometry showed DNA content beyond diploid, which could suggest the activation of additional ploidy maintenance checkpoints at diploidization or that higher ploidy is not viable in the mutants.

A third tempting explanation for the ploidy phenotype in *scp160Δ* cells is nondisjunction events due to aberrant chromatin structure, given the role of Scp160p/vigilin in heterochromatin formation and silencing of telomeres and the mating locus (see Review Article). However, this is scenario is unlikely as the silencing of telomeric regions was shown to be independent of ploidy in *scp160Δ* cells (Marsellach et al., 2006). To assay telomeric silencing, Marsellach and colleagues inserted the *URA3* gene (encoding an enzyme that catalyses a step in the uracil synthesis pathway) into the normally silenced telomeric region. When silencing is lost and *URA3* is expressed, the resulting enzyme leads to cell death in the presence of 5-fluoroorotic acid (5-FOA) as Ura3p catalyses the compound into the toxic 5-fluorouracil (Boeke et al., 1987). Performing such an experiment with the Scp160p(KH1-10) truncation would allow further mapping of the KH domains which are involved in mediating heterochromatin formation and silencing. Moreover, chromatin-IP (ChIP) analysis showed that deposition of the silencing protein Sir3p at telomeric regions is reduced in *scp160Δ* cells (Marsellach et al., 2006), and it would be of interest to investigate whether this also occurs in a strain expressing the Scp160p(KH1-10) truncation. Strikingly, this study in *S. cerevisiae* contrasts observations in *Drosophila* S2 cells in which heterochromatin was destabilized when KH13-14 of the *Drosophila* vigilin homolog DDP1 was expressed in addition to the full-length protein (Zhou et al., 2008). The authors of the *Drosophila* study reasoned that KH13-14 are involved in mediating heterochromatin formation and that its ectopic expression sequesters important factors in this process. However, the authors did not test whether heterochromatin formation was lost with a DDP1(KH1-12) truncation and ectopic expression of a KH10-14 fragment did not show heterochromatin destabilization (Zhou et al., 2008). Alternatively, this may reflect a difference in the structure-function coordination of the various vigilin homologs. In support, loss of the

S. pombe vigilin homolog Vgl1 had no effect on silencing at telomeres and the mating-type locus as assayed by the *URA3* method described above.

4.1.2 – Separation of Scp160p and Bfr1p function in polyQ toxicity and biosynthesis

In contrast to ploidy maintenance, analyses of the polyQ reporters showed a separation of Scp160p and Bfr1p function. Whereas loss of Scp160p suppressed the toxicity of the 97Q(CAG)mCh reporter, loss of Bfr1p instead resulted in severe toxicity from all three polyQ reporters, including 97QmCh and 104Q(CAA)mCh which were only slightly toxic in wild-type cells. This observation would suggest that the toxicity conferred to *scp160Δ* cells is primarily dependent on codon usage and may occur in a translation elongation-related manner (see section 4.3). In contrast, the toxic effect of the polyQ reporters in *bfr1Δ* cells is likely to depend primarily on the amino acid sequence/protein. As Bfr1p has been implicated in ER-to-Golgi and intra-Golgi transport (Jackson and Képès, 1994), the toxicity of the polyQ reporter proteins may result from perturbed vacuolar transport and the secretory pathway. Indeed, expression of expanded polyQ and other aggregation-prone reporter proteins in *S. cerevisiae* have been reported to induce ER stress by overloading the protein-folding capacity of the ER (Duennwald and Lindquist, 2008; Leitman et al., 2013; Low et al., 2014), which activates the UPR pathways (Travers et al., 2000). One consequence of UPR activation is the Ire1p-mediated up-regulation of a number of genes of the secretory pathway, which is believed to alleviate ER load by promoting anterograde transport of secreted and membrane proteins (Kimata et al., 2006; Travers et al., 2000; Walter and Ron, 2011). Notably, *BFR1* is among the genes up-regulated upon activation of the UPR (Kimata et al., 2006; Travers et al., 2000), and Bfr1p levels are elevated upon elevated ER stress (Low et al., 2014).

Strikingly, the toxicity of the polyQ reporter proteins to *bfr1Δ* cells is in contrast to two genomic screens in which the *bfr1Δ* was found to suppress polyQ-induced toxicity (Giorgini et al., 2005; Kaiser et al., 2013). This discrepancy could lie in the genetic background of the yeasts used in the screens (BY4741, as part of the yeast deletion library) and this study (W303a). In support of this possibility, Serpionov et al. reported recently that expanded polyQ induces toxicity differently among two different *S. cerevisiae* backgrounds (BY4742, differing from BY4741 in mating type vs 74-D694). Additionally, differences in experimental approach (high-throughput screen vs

serial dilution assay) and construction of the polyQ reporter can also influence the toxicity (Duennwald et al., 2006; Hofer et al., 2018). Consistent with this scenario, the reporter used in the screen by Giorgini and colleagues carry a FLAG epitope, whose presence anywhere on a polyQ reporter (even for one containing 25Q's, a number below the pathogenic threshold in humans) has been shown to confer toxicity not observed with an HA epitope (Duennwald et al., 2006). Similarly, one of the polyQ reporters used by Kaiser and colleagues consisted only of 56Q's and a C-terminal YFP and the other reporter was the same one used by Giorgini and colleagues.

The difference in polyQ-induced toxicity between *scp160Δ* and *bfr1Δ* cells suggest that the effect occurs at various stages of protein synthesis and lifespan. While the conserved nature of Scp160p/vigilin is exciting in the context of Huntington's and other polyQ-expansion diseases (see also section 4.4), the restriction of Bfr1p to *Ascomycetes* means its role may be fungal-specific. Alternatively, given the role of Bfr1p in anterograde vesicle transport and its link to the UPR, these processes and pathways represent an interesting and on-going area of research in polyQ-expansion diseases.

4.2 – Scp160p may impact many cellular processes via Q/N-rich-mediated protein aggregation (related to Cheng et al., 2018)

Scp160p, the *S. cerevisiae* homolog of vigilin, was previously reported by our lab to enhance translation elongation efficiency of many mRNA in a manner dependent on codon correlation (see section 1.4 and Hirschmann et al., 2014). In Cheng et al., 2018 (full-text in Appendix), I investigated this function further using three polyQ reporters that differ in the codon usage of their polyQ-encoding region. Strikingly, while all three reporters encode the same protein sequence based on exon 1 of mHtt (the additional 7 Q's in 104Q(CAA)mCh had a negligible impact; Cheng et al., 2018), codon usage was a major determinant in polyQ-induced toxicity. Furthermore, as all three reporter proteins aggregated in wild-type cells, polyQ-induced toxicity is separate from aggregation. It is interesting to note that the severely toxic variant – 97Q(CAG)mCh – also accurately constitutes mHtt at the nucleic acid level. Moreover, the toxicity conferred by 97Q(CAG)mCh required an in-frame start codon upstream of the polyQ-encoding (CAG-repeat) sequence and was suppressed by the loss of Scp160p. Together, these results argue that in our experimental system, polyQ-induced toxicity occurs at the level of translation of the mRNA. However, the mechanism(s) through

which 97Q(CAG)mCh confers toxicity in wild-type cells, and how this toxicity is suppressed in *scp160Δ* cells remains to be investigated. Unfortunately, analyses of this reporter are made difficult by the low abundance of 97Q(CAG)mCh mRNA and protein in wild-type cells. For example, it is as-yet unknown if this reflects increased mRNA degradation or poorer transcription in wild-type cells compared to *scp160Δ* cells. Yet, the preliminary RT-qPCR data on 97Q(CAG)mCh mRNA level in *dom34Δ* cells suggests that the high abundance of the mRNA in *scp160Δ* cells results from more than just escape from the RQC. Comparing the half-life of 97Q(CAG)mCh in wild-type and *scp160Δ* cells will be one step toward answering this question. In parallel, the abundance of 97Q(CAG)mCh mRNA may be assessed in *xrn1Δ* or *skiΔ* cells to probe whether the mRNA is targeted by these RNA decay machineries. Moreover, identification of factors that bind 97Q(CAG)mCh in wild-type and *scp160Δ* cells (in an *xrn1Δ* or a *skiΔ* background to stabilize mRNAs (Parker, 2012)) would highlight the potential changes in regulation of the mRNA upon loss of Scp160p.

A second striking observation of the polyQ reporter proteins in this work is their reduced aggregation in *scp160Δ* cells. Importantly, this effect was independent of the reporters' codon usage, suggesting that Scp160p's role in polyQ-mediated aggregation is not related to its role in codon-related translation. A combined filter-trap binding and quantitative mass spectrometry approach also revealed that the aggregation of many endogenous Q/N-rich proteins is reduced in *scp160Δ* cells under normal growth conditions. Although the codon usage of the proteins identified by our approach has not been analyzed, the fact that a mixture of both glutamine-encoding codons – CAA and CAG – encodes polyQ regions of many proteins in *S. cerevisiae* (Mar Albà et al., 1999) is consistent with the idea that Scp160p's influence on Q/N-mediated protein aggregation is independent of its role in codon-related translation. Analysis of the "cellular component" Gene Ontology (GO) terms associated with the Q/N-rich proteins with reduced aggregation in *scp160Δ* cells showed enrichment for "Nuclear pore central transport channel" (p-value=5.16E-05), "Cytoplasmic stress granule" (p-value=4.04E-06), and "Mating projection tip" (p-value=1.90E-04) (PANTHER Classification System, Mi et al., 2017). Interestingly, these GO terms aligns with some of Scp160p's proposed role in P-body formation (Weidner et al., 2014) and mating (Aronov et al., 2007; Gelin-Licht et al., 2012). Most notably, the number and functional variation (*Saccharomyces* Genome Database) of the affected

Q/N-rich proteins demonstrates a potential for Scp160p to influence many processes and protein interaction networks in this manner.

Two of the proteins from the mass spectrometry approach which we validated were Cyc8p and Nab3p, whose polyQ regions have been reported to be biologically important (Figure 4.2 and Gemayel et al., 2015; O'Rourke and Reines, 2016). Cyc8p is a conserved transcription repressor complex regulating >150 *S. cerevisiae* genes involved in flocculation, glucose repression, mating type, and sporulation (Gemayel et al., 2015; Smith and Johnson, 2000; Varanasi et al., 1996). It's regulatory activity is affected by the protein's propensity to aggregate, which in turn is determined by the length of the QA-repeat and polyQ-regions (together named TR2) in the middle of the protein (Gemayel et al., 2015). Notably, the expression of 27 genes changed in response to TR2 length in glucose-rich media (by RNA-Seq; Gemayel et al., 2015), of which 6 were observed to be downregulated and 1 to be upregulated in response to depletion of Scp160p in a previous unpublished study (by microarray; Schreck, 2010).
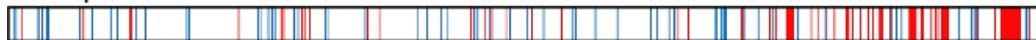
Figure 4.2 – Schematic of the Q/N-rich proteins Cyc8p and Nab3p. Glutamine (Q) residues are represented in red and asparagine (N) in blue.

Nab3p is an RBP involved in the transcription termination and 3' processing of a variety of RNA species including mRNA, cryptic unstable transcripts, small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA) (Porrua and Libri, 2015; Webb et al., 2014). *In vitro* studies have shown that a C-terminal fragment of Nab3p containing a polyQ-region/LCR can form hydrogels and self-assemble into amyloids, and that this region was required for viability *in vivo*. Specifically, it was the polyQ-richness of this fragment which was important, as substitution of every second glutamine in this region to glutamic acid also disrupted amyloid formation *in vitro* and lead to inviability *in vivo* (Loya et al., 2013; O'Rourke and Reines, 2016).

The result of the mass spectrometry approach presents a tempting explanation to reconcile the many phenotypes associated with *scp160Δ*, that Scp160p can facilitate the Q/N-mediated aggregation important for the biological function of many

proteins. Given the prevalence of polyQ- and LCR-regions in transcription factors and RBPs, and the role of aggregation in their function, this represents the potential to have widespread downstream effects on gene expression and cell function. However, further work will be required to investigate how Scp160p facilitates the aggregation of these proteins (see section 4.3) and what determines Scp160p's specificity towards them.

4.3 – Possible mechanism(s) of Scp160p function in protein aggregation

4.3.1 – Altered protein quality control

A major finding of this work is the role of Scp160p in polyQ- and Q/N-mediated protein aggregation, an ongoing direction of study stemming from this work will be the mechanism(s) behind this involvement. As this work first assayed of three overexpressed aggregation-prone reporter variants, I asked whether the effect might have stemmed from changes in protein quality control. Molecular chaperones are major factors in maintaining protein quality and have the ability to modulate aggregation of prions, expanded polyQ-, and misfolded proteins (Ciechanover and Kwon, 2017; Duennwald et al., 2012; Sakahira et al., 2002; Walters et al., 2015). One possibility is that the loss of Scp160p altered chaperone expression and thereby reduced polyQ- and Q/N-mediated protein aggregation. Indeed, unpublished work from our lab has shown by microarray that depletion of Scp160p resulted in the downregulation of several HSP genes (Schreck, 2010). However, the work presented here does not support this scenario since: 1) a preliminary screen of select molecular chaperones by western blot showed no significant changes in protein abundance between wild-type and *scp160Δ* cells (Figure 3.6.1A), 2) quantitative mass spectrometry of the WCL showed negligible changes in relative protein abundance of these chaperones, and 3) combining an *hsp31Δ* (the HSP with the greatest increase in protein abundance) with an *scp160Δ* did not restore 97QmCh aggregation (Figure 3.6.1B). Interestingly, the quantitative mass spectrometry of the WCL (this work) showed that the mitochondrial-matrix chaperone Hsp78p was 1.51 times more abundant in *scp160Δ* cells compared to wild-type (although only detected in one of two mass spectrometry replicates). However, this is in contrast to the microarray observation that *HSP78* transcript levels were decreased upon Scp160p-depletion to 0.31-fold that of a no depletion (Schreck, 2010).

The DJ-1 superfamily member Hsp31p has been shown to deter [PSI+] prion formation as well as α-synuclein and polyQ reporter aggregation. Therefore, it was a promising candidate to test as an effector of reduced 97QmCh aggregation downstream of the *scp160Δ*. The failure of an *hsp31Δscp160Δ* strain to restore 97QmCh aggregation suggests that reduced aggregation in *scp160Δ* does not function through Hsp31p. However, it cannot be ruled out that loss of Hsp31p in an *scp160Δ* background results in compensatory upregulation of other chaperone members, especially those of the DJ-1 family. Indeed, Hsp31p is one of four DJ-1 homologs in *S. cerevisiae*, the others being Hsp32p, Hsp33p, and Hsp34p (Aslam and Hazbun, 2016). This scenario could be tested by RT-qPCR to assess if *HSP32*, *HSP33*, and/or *HSP34* expression are increased in *hsp31Δ* and *hsp31Δscp160Δ* cells. Alternatively, 97QmCh aggregation could be scored in strains carrying the *scp160Δ* in combination with *hsp31Δ* and other DJ-1 family members. It should be noted that due to a 99% identity between Hsp32p, Hsp33p, and Hsp34p (Wilson, 2014), assessing their protein abundance by western blot could prove difficult.

In addition to the action of molecular chaperones such as the HSPs, misfolded and aggregated proteins can be targeted to three spatially separated quality control compartments: the intra-nuclear quality control compartment (INQ, formerly juxtanuclear quality control compartment JUNQ), the insoluble protein deposit (IPOD), and the cytosolic quality control compartment (CytoQ) (Kaganovich et al., 2008; Miller et al., 2015; Mogk and Bukau, 2017). These quality control compartments allow refolding or degradation of multiple misfolded proteins at once. Unlike the INQ and CytoQ, which are induced under stresses like heat-shock, the IPOD is a destination for misfolded amyloid proteins under non-stress conditions, where they are degraded via autophagy (Kaganovich et al., 2008; Miller et al., 2015). Interestingly, aggregated forms of mHtt-based reporters were deposited in or adjacent to the IPOD but not the INQ (Kaganovich et al., 2008; Yang et al., 2016). Thus, it could be that the higher aggregation of polyQ reporters in wild-type cells observed in this work represent more efficient recognition and sequestration of these proteins into the IPOD. However, this scenario does not reconcile with the suppression of 97Q(CAG)mCh-mediated toxicity in *scp160Δ* cells. Alternatively, it may be that IPOD-sequestered polyQ reporters are more efficiently degraded in *scp160Δ* cells, resulting in less observed aggregates at steady state. To distinguish between these two scenarios, 97QmCh aggregates can be assayed for colocalization with a fluorescently-tagged IPOD markers such as

Rnq1p (Kaganovich et al., 2008; Kayatekin et al., 2014). Greater colocalization of 97QmCh with Rnq1p would argue for more efficient IPOD-targeting in wild-type cells. Another assay that would address these scenarios is to induce 97QmCh protein expression and follow its stability after repressing its expression by addition of glucose to the culture medium. If 97QmCh is less stable in *scp160Δ* cells than wild-type cells, this would support the scenario where IPOD-sequestered polyQ proteins are degraded more efficiently in the deletion.

Finally, ribosome-associated quality control is another means to ensure protein quality in the cell. Part of this work asked whether translation elongation in *scp160Δ* cells may be exacerbated by the biosynthesis of the aggregation-prone 97QmCh reporter protein and trigger RQC. Strikingly, 97QmCh protein accumulation and aggregation were further reduced in *ltn1Δscp160Δ* cells compared to *scp160Δ* cells. Western blot analysis of 97QmCh probed against its N-terminal myc-epitope showed that the protein runs slightly smaller in *ltn1Δ* cells compared to wild-type, *scp160Δ*, and even *ltn1Δscp160Δ* cells. Given the role of Ltn1p in poly-ubiquitination of stalled nascent peptides, it is tempting to reason the downshift of 97QmCh is due to the loss of Ltn1p-mediated ubiquitination in *ltn1Δ* cells. However, this downshift is not observed in *ltn1Δscp160Δ* cells. Nevertheless, the ubiquitylation-status of 97QmCh in wild-type, *scp160Δ*, *ltn1Δ*, and *ltn1Δscp160Δ* cells may be assessed by IP of 97QmCh (via its N-terminal myc-epitope) followed by immunoblotting against ubiquitin. In parallel, treatment of lysates with a recombinant deubiquitintating enzyme Usp2 would also allow assessment of ubiquitylation status (Yonashiro et al., 2016).

4.3.2 – Aberrant translation elongation kinetics

An alternative mechanism to explain the reduced polyQ- and Q/N-mediated protein aggregation in *scp160Δ* cells would be altered translation elongation kinetics. As described in section 1.3, the elongation kinetics of a nascent peptide can have a great impact on the co-translational folding, ultimate conformation, and functionality of the protein product. In support, slowing down elongation with a mild dose of translation inhibitors or growth at low temperatures have been shown to influence protein folding and to promote native folds and protein function (Meriin et al., 2012; Sherman and Qian, 2013; Zhou et al., 2013). Furthermore, it has been speculated for wild-type unexpanded huntington protein that translation of the proline-rich region (PRR), which occurs slowly, promotes co-translational binding of chaperones to the N-terminal 17

amino acids which facilitate proper folding and processing (Nissley and O'Brien, 2016). Importantly, the PRR – and thus elongation slowdown – should be spaced less than 35 amino acids (the threshold number of Q's before pathogenesis) so as to facilitate this process as the N-terminal 17 amino acids emerge from the ribosome exit tunnel. Given Scp160p's role in enhancing translation efficiency, it is possible that the absence of Scp160p slows elongation of the polyQ reporters and thereby facilitate their folding and/or processing to a less aggregation-prone conformation. It is notable that aggregation of the polyQ reporters are not abolished – only reduced – in *scp160Δ* cells, which may reflect the inherent nature of polyQ regions to form coiled-coils that facilitate protein-protein interaction and their spontaneity to aggregate (Fiumara et al., 2010; Schaefer et al., 2012).

In support of this, treating *S. cerevisiae* cultures with a mild dose of CHX (30ng/µL) reduced aggregation – but not abundance – of the 97QmCh reporter protein (see section 3.7). However, it will have to be investigated in greater detail whether loss of Scp160p indeed reduces polyQ aggregation by slowing down translation elongation. Attempts to assess Cyc8p-GFP aggregation by mild CHX were inconclusive, likely owing to the relative low abundance of the endogenous protein (*Saccharomyces* Genome Database: https://www.yeastgenome.org/locus/S000000316/protein). Further optimization of CHX dosage will be required to better assess its effect on Cyc8p-GFP. As an alternative to mild CHX treatment, it may also be possible to slow elongation by growing *S. cerevisiae* cultures at a lower temperature (e.g. 18˚C). In parallel to such translation elongation slowdown assays, it would be prudent to assess how the fold-conformation of 97QmCh and Cyc8p-GFP changes between wild-type and *scp160Δ* cells, as well as in the presence and absence of elongation slowdown. This could be tested by assaying sensitivity to proteases, whose access to cleavage sites would be affected by the target protein's fold and conformation. Differences in protease sensitivity would add support to the hypothesis that elongation slowdown in the absence of Scp160p can shift the fold of polyQ-containing and Q/N-rich proteins to ones that are less aggregation-prone. However, when considering such an assay, one should separate the soluble polyQ-containing and Q/N-rich proteins from the aggregated forms, which may be done by differential centrifugation (Meriin et al., 2003).

Finally, the reduction of polyQ reporter aggregation in *scp160Δ* cells may reflect a difference in the conformation of the aggregates itself. To this end, the conformation of 97QmCh, 97Q(CAG)mCh, and 104Q(CAA)mCh aggregates in wild-type and *scp160Δ* cells can be compared by CLEM. This technique allows the observation of ultra-structures of fluorescently-labeled entities (de Boer et al., 2015). Observation of GFP-tagged polyQ reporters in HeLa and *S. cerevisiae* cells revealed that the aggregates can adopt different ultra-structures that correlate with their dynamic properties (Peskett et al., 2018). CLEM comparison of polyQ aggregates in wild-type and *scp160Δ* cells can thus also be complemented by testing the molecular dynamics within the aggregates via fluorescence recovery after photobleaching (FRAP).

In addition to probing protein fold and aggregate conformation, it would also be important to assay elongation rate of the polyQ reporters and Q/N-rich proteins. For example, elongation rate of 97QmCh in wild-type and *scp160Δ* cells may be compared by metabolic labeling. In this case, *de novo* 97QmCh are labeled and followed at specified time points using a pulse of radioactive [^{35}S]-methionine. A comparison of the amount of [^{35}S]-methionine incorporated into 97QmCh in wild-type and *scp160Δ* cells at each time point can provide a measure for the rate of translation. Such a method will require optimization of both the length of the [^{35}S]-methionine pulse as well as the efficiency and specificity of immunoprecipitation for the protein of interest. These optimization steps are currently underway (by Jonathan Feicht in collaboration with Dr. Tobias Jores and Prof. Dr. Doron Rapaport). If [^{35}S]-methionine incorporation into 97QmCh occurs slower in *scp160Δ* cells, this would support the hypothesis that polyQ-mediated aggregation is negatively influenced by perturbed elongation kinetics in the absence of Scp160p. In such a scenario, the elongation rates of the other two polyQ reporters (97Q(CAG)mCh and 104Q(CAA)mCh) and/or endogenous Q/N-rich proteins may also be assayed.

Another powerful method to probe genome-wide *in vivo* elongation rates, with nucleotide-resolution, is ribosome profiling (Ingolia et al., 2009). This technique exploits the protection of ribosome-occupied mRNA fragments from ribonucleases, which are then deep-sequenced. Alignment of the sequenced mRNA fragments provides a snapshot of the distribution of ribosomes over the transcriptome. Since the time spent by ribosomes at a codon or region relates to ribosome transit speed, ribosome footprint densities can also reveal co-translational events such as ribosome slowdown or stalling (Ingolia et al., 2012). Comparing the ribosome footprint densities

for the same gene in two conditions or strains can provide information on changes in elongation speeds. Therefore, ribosome profiling is an ideal method to investigate how loss of Scp160p can influence elongation rate of the transcriptome, and more importantly, if there are certain genes and/or sequences (e.g. LCRs or Q/N-rich regions) which are more affected by the loss.

By combining ribosome profiling with the mass spectrometry data in this work, it would be possible to compare the elongation kinetics of those specific Q/N-rich genes whose aggregation is reduced in *scp160Δ* cells. If ribosome footprint densities are higher in the *scp160Δ* translatome at corresponding regions within transcripts of interest, it would argue that elongation efficiency is perturbed in the absence of Scp160p. Such an observation would also be consistent with the hypothesis that lowered elongation rate in *scp160Δ* cells can promote co-translational folding and adoption of less aggregation-prone conformations by Q/N-rich proteins. If ribosome transit is slowed for Q/N-rich protein-encoding genes in *scp160Δ* cells, it will be of interest to analyze at which regions/sequences this occurs. Does loss of Scp160p result in slowdown of translation over the entire transcript, at Q/N-rich or LCR-encoding regions, or the region(s) adjacent? Such information would shed light on how elongation kinetics, and Scp160p, affects co-translational folding and protein biosynthesis.

Recently, selective ribosome profiling has also been performed in which addition of an immunoprecipitation step allows a view of a subset of ribosomes and/or ribosome-nascent chain complexes. Such a method has been employed to study binding of chaperones to nascent chains, as well as to determine the order of co-translational protein complex assembly (Becker et al., 2013; Shiber et al., 2018). This method may also be used to study Scp160p-associated ribosomes and provide information on the transcripts – and sequences within – that are specifically regulated by Scp160p. Scp160p-specific ribosome profiling would be a great complement to the analysis of the whole wild-type and *scp160Δ* translatomes.

Provided here are possible scenarios to link Scp160p with polyQ- and Q/N-mediated protein aggregation. While these avenues are worth pursuing, the previously published involvement of Scp160p in enhancing translation efficiency makes the elongation kinetics and co-translational folding hypothesis a particularly attractive one to test. Due to the importance of polyQ and Q/N-rich regions in protein aggregation, it will be important to investigate how Scp160p affects this process. Our understanding

of this can identify Scp160p has an important factor in protein biosynthesis. Finally, given the variety of functions associated with Q/N-rich proteins, this area of research can shed light on the distinct phenotypes associated with loss of Scp160p.

4.4 – Implications of Scp160p/Vigilin in Q/N-rich protein aggregation for higher organisms

The work presented here have shown a role for the *S. cerevisiae* homolog of the vigilin family of RBP, Scp160p, in facilitating aggregation of polyQ- and Q/N-rich proteins. Given the highly conserved nature of the vigilin proteins, this work could have implications in higher eukaryotes. In particular, polyQ- and Q/N-mediated protein aggregation are important areas of research in higher eurkaryotes due to their link to both neurodegenerative diseases as well as mediation of normal protein function and interactions (see section 1.6). The polyQ reporters used in this study were derived from the pathogenic expanded huntingtin protein, which causes HD in humans (Fan et al., 2014), thereby making vigilin an interesting target for future research in HD pathogenesis. However, care must be taken when undertaking such translationary studies due to differences between *S. cerevisiae* and mammalian systems. Perhaps the biggest factor to consider is the endogenous context of the *HTT* gene. *S. cerevisiae* is a great isolated system in which to study mHtt protein aggregation and toxicity, due to the absence of *HTT* gene. However, wild-type *HTT* is endogenous to mammalian systems and plays an important role, which adds a layer of complexity to understanding mHtt aggregation and pathogenesis. Therefore, although Scp160p acts on the mHtt-derived reporter, this may not be representative of vigilin and mHtt in a mammalian context where mHtt biosynthesis and pathogenicity may depend on other mammalian-specific factors. Accordingly, the ability of the expanded CAG trinucleotide repeat in m*HTT* to self-complement into a stem has been reported to contribute to pathogenesis by sequestering the splicing factor muscleblind-like 1 (MBNL1) into nuclear foci in human fibroblasts derived from HD patients (de Mezer et al., 2011). Although the existence of a similar factor that binds CAG-stems cannot be ruled out, no *S. cerevisiae* homolog of MBNL1 can be found by homology search.

In addition to the mHtt-derived polyQ reporters, the present work also showed the influence of Scp160p on the aggregation of many endogenous unexpanded Q/N-rich proteins in a native non-disease context. As emerging research demonstrate the biological importance of Q/N-mediated aggregation and interaction on protein function,

the potential impact of vigilin on Q/N-mediated protein aggregation in higher eukaryotes will be an interesting aspect to study. Notably, the vigilin-interaction partner heterochromatin protein 1 (HP1) (Cheng and Jansen, 2017) has been shown to mediate heterochromatin formation in *Drosophila* and humans by undergoing phase separation (Larson et al., 2017; Strom et al., 2017).

The highly conserved nature of the vigilin RBPs make translational studies in higher eukaryotes exciting, both in the context of repeat expansion diseases and in normal Q/N-rich protein function. However, as alluded to earlier, differences between the unicellular *S. cerevisiae* and multicellular mammalian systems is a major caveat to consider. In the case of protein biosynthesis and pathogenesis in repeat expansion diseases, the proper cell type must be used. Although huntingtin and vigilin proteins are reportedly expressed in many cell types (Hofer et al., 2018; Uhlén et al., 2015), the cells used should reflect and allow assay of processes affected in neurodegeneration. Moreover, given the codon context of Scp160p's role in translation, the difference in codon usage profile between *S. cerevisiae* and mammals might also make an impact. For example, in *S. cerevisiae* CAG is the infrequent glutamine-encoding codon and CAA is the frequent codon, which is opposite to that in mammals (Athey et al., 2017; Sharp and Li, 1987). Despite these differences, many studies have demonstrated a level of transferability and mechanistic conservation from *S. cerevisiae* to mammalian cells in protein aggregation and pathogenesis associated with repeat expansion diseases. The potential influence of vigilin in disease-related polyQ-, as well as normal Q/N-mediated protein aggregation in mammalian systems will be an interesting avenue for future research.

5. MATERIALS AND EXPERIMENTAL PROCEDURES

5.1 – Oligonucleotides used in this study

Name	Sequence (5' to 3')	Database Number (RJO)
scp160_ko_HIS3MX6 fw S1	TAAAATATACTTCCCACACCCCCTCCTTCCATTAT AACTGCACGTACGCTGCAGGTCGAC	2509
scp160_ko_HIS3MX6 rev S2	GCCAAAATCTATATTGAAAAAAATTGGTTTCAAAG AGCTTGTATCGATGAATTCGAGCTC	2510
SCP160 3'-UTR rev	GTTCGTGTTACTTCAACTTCA	2511
Scp160_ko_check_1	GCTTCAGCAGTTCCTTGTTCA	2581
Scp160_ko_check_2	GCAGCTACACCAGAAACCAAAAG	2582
BamHI_Scp160_1	CGCAAAGGATCCATGTCTGAAGAACAAACCGC	2583
Scp160_PstI_2	CGCAAACTGCAGCTATCTTCTTAAGGATTTCAAAA CC	2584
Scp160_seq1	TCCTCGGAAGAAGTTGGTGC	2665
Scp160_seq2	CGGAGTCTGCGTTAGGAATATC	2666
Scp160_seq3	AAGTACCGTGTGGACCAATC	2667
Scp160_seq4	AGGTGGCCTTCTTGTTCAG	2668
Scp160_seq5	AGATTTTTGGCATGGGTCAG	2671
Scp160_seq6	CACCCTCATCGGAGGCAAAC	2672
Scp160_seq7	TCACGAATAGTTCTACCACCTG	2673
Scp160_ORF_EcoRI_rev	GCTCGATGGAGAATTCAAAATAG	2702
Scp160_ORF_EcoRI_for	CTATTTTGAATTCTCCATCGAGC	2703
Act1_qPCR_1_for	TCAGAGCCCCAGAAGCTTTG	2920
Act1_qPCR_1_rev	TTGGTCAATACCGGCAGATTC	2921
PGK1-RT_F	GAACGGTCCACCAGGTGTT	4132
PGK1-RT_R	GACGGTGTTACCAGCAGCAG	4133
18S_F	TCAACACGGGGAAACTCACC	4139
18S_R	CTAAGAACGGCCATGCACCA	4148
dBFR1_S1	AACGTAATAGCATATTTTCTAACAACACAGCCATT GCCATGCGTACGCTGCAGGTCGA	4853
dBFR1_S2	GAAGAAAGATCAGGAGAAAAATTTTTTTCTACTTC AGGTTTAATCGATGAATTCGAGCTCG	4854
Bfr1_cds_f	GTGATGATGTCAAGATCACCG	4875
Bfr1_3utr_r	TCATCGTCACACCCTATTGAC	4876
His3_cas_f	AGCACGAAGGGAGTGTTGTAA	4877
bfr1_genomic_del_f	TAACTGATCTCGACGACGTTG	4918
bfr1_genomic_del_r	GAGGAAAGAATTGGCTGGTAAG	4919
N-term mCherry fw	GAGTTCATGCGCTTCAAGGT	4926
N-term mCherry rv	GTCTGGGTGCCCTCGTAG	4927
dom34_S1	CGTTGTCATTTTGTTCAATTATCGCATTCCTATCA TAGCAAAAATCGTACGCTGCAGGTCGA	5120
dom34_S2	GTTGCAAATTTTATGTGTACATTACTTTTTTCTTAC ATAGTAAATATCGATGAATTCGAGCTCG	5121

dom34del_chk_f	GATGGACTCCGTAGAGAGCAAT	5122
dom34del_chk_r	TGCACGTGACTAACTAGAACC	5123
cyc8tag_fw	GAAAATGTAGTAAGGCAAGTGGAAGAAGATGAAAACTACGACGACCGTACGCTGCAGGTCGA	5153
cyc8tag_rv	TCGTTGATTATAAATTAGTAGATTAATTTTTTGAATGCAAACTTTATCGATGAATTCGAGCTCG	5154
cyc83utr_rv	GGATTCAACAAGGGAAACCG	5155
scp160del_kh1to6_f	TCTTTGAACAAGGCCAATGAGTCATTAAACTCTCTACGTACTAAACGTACGCTGCAGGTCGAC	5346
scp160del_kh1to7_f	AAAGCTGCTAACAAAATTTTTGAATCTATTTTGAATTCTCCATCGCGTACGCTGCAGGTCGAC	5347
scp160del_kh1to8_f	AACTTGACTCATGCCAAGAAATATTTGGCTGCCGAAGCCAAAAAACGTACGCTGCAGGTCGAC	5348
scp160del_kh1to9_f	AATAAAGCTCATGAAGAATTGAAAGCTCTGTTAGATTTTGAAATGCGTACGCTGCAGGTCGAC	5366
scp160del_kh1to10_f	AACATAAAAGATGCCGCCAAACGTGTGGAATCCATTGTTGCCGAAGCCCGTACGCTGCAGGTCGAC	5367
scp160del_kh1to11_f	GTCAAAAAGGTTGTCGAGGAAATTAACAAGATTGTCAAGGATGCTCGTACGCTGCAGGTCGAC	5368
TR2check_f	TCAATGGTACAACAACAGCATCCTGCTCAA	6213
TR2check_r	TATGGTTGCCCTTGTTGAGGATTTAACATT	6214
TR1check_f	AGACTAGTACTACAACTACAACAGCA	6215
TR1check_r	GTTTCTGCCAAAGAAGCAAT	6216
ltn1_del_f	ATCTGCTAAGCCATCAAAAAAAGTTCAAGCAATAGTTGGTTCTTACGTACGCTGCAGGTCGAC	6273
ltn1_del_r	GTTTAAAAAATGTAGTACATTTATATGAAATTTATATGCGATAGTATCGATGAATTCGAGCTCG	6274
ltn1_5UTR_f	CAATACTGAAGAAGTCCTTCTTA	6275
ltn1_3UTR_r	GGTATAGGGCTGGATTGTATAA	6276
ltn1_orf_f	GCCTACGAGCCTAGCTTTAGCACC	6277
Nab3_Q1_f	ACAAACTATTACCAGGGTTACAGT	6457
Nab3_Q1_r	TGGTGGCATCCCATAATTACC	6458
Nab3_Q2_f	ACGTTGTATCGAATTTGCTTTCA	6459
Nab3_Q2_r	AGGATGAGTTCATAGAGGAATATCC	6460
Nab3_Q3_f	GGTTTAATACAATCAATGCAAGGC	6461
Nab3_Q3_r	GACTTTGAACATTATTGCCAGC	6462

5.2 – Plasmids used in this study

Name	Description	Database Number (RJP)	Reference
pFA6a-HIS3MX6	-	135	(Longtine et al., 1998)
pYM3	6xHA::klTRP1	277	
pYM5	3xmyc::HIS3MX6	279	(Knop et al., 1999)

pRS313	pRS313::HIS3MX6	721	-
pYM44	yeGFP::HIS3MX6	1243	(Janke et al., 2004)
pFA6a-KANMX6	-	1347	(Longtine et al., 1998)
Tet-off-*SCP160*	pCM182-*SCP160*::klTRP1	1463	(Hirschmann et al., 2014)
97QmCh	pYES2-97QmCh::URA3	1920	(Park et al., 2013)
97Q(CAG)	pMS-RQ-97Q-CAG	1935	GeneArt AG
104Q(CAA)	pMS-RQ-104Q-CAA	1936	GeneArt AG
97Q(CAG)mCh	pYES2-97Q(CAG)mCh::URA3	1937	This study
104Q(CAA)mCh	pYES2-104Q(CAA)mCh::URA3	1938	This study
(-)ATG-97Q(CAG)mCh	pYES2-(-)ATG-97Q(CAG)mCh::URA3	2055	This study
92Q(CAA)mCh	pYES2-92Q(CAA)mCh::URA3	2163	This study
92Q(CAA)	pMS-RQ-92Q-CAA	2164	GeneArt AG
pRS313-*SCP160*	pRS313-*SCP160*::HIS3MX6	2179	This study

5.3 – Yeast strains used in this study

Name	Relevant Genotype	Reference
W303a	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	-
W303a/α	*MATa/MATα leu2-3,112/leu2-3,112 trp1-1/trp1-1 can1-100/can1-100 ura3-1/ura3-1 ade2-1/ade2-1 his3-11,15/his3-11,15*	-
RPY497	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1*	(Hirschmann et al., 2014)
RPY3178	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::HIS3MX6*	(Hirschmann et al., 2014)

RPY3925	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160-GFP::HIS3MX6*	This study
RPY4584	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-97QmCh::URA3*	This study
RPY4588	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97QmCh::URA3*	This study
RPY4709	*MATa/MATα leu2-3,112/leu2-3,112 trp1-1/trp1-1 can1-100/can1-100 ura3-1/ura3-1 ade2-1/ade2-1 his3-11,15/his3-11,15 scp160::HIS3MX6/SCP160 bfr1::klTRP1/BFR1*	This study
RPY4749	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-97Q(CAG)mCh::URA3*	This study
RPY4750	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97Q(CAG)mCh::URA3*	This study
RPY4754	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-104Q(CAA)mCh::URA3*	This study
RPY4755	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-104Q(CAA)mCh::URA3*	This study
RPY4778	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::HIS3MX6 bfr1::klTRP1*	This study
RPY4814	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::HIS3MX6 pTet-off-SCP160::klTRP1 pYES-97QmCh::URA3*	This study
RPY4816	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::HIS3MX6*	This study
RPY4825	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 bfr1::HIS3MX6 pYES-97QmCh::URA3*	This study
RPY4867	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 CYC8-GFP::HIS3MX6*	This study
RPY4874	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 CYC8-GFP::HIS3MX6*	This study
RPY4890	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160-GFP::HIS3MX6 pYES-97QmCh::URA3*	This study

RPY4896	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97QmCh::URA3 pUG34-GFP-SCP160*	This study
RPY4916	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-12-GFP::HIS3MX6*	This study
RPY4919	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-12-GFP::HIS3MX6 pYES-97QmCh*	This study
RPY4960	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-6-GFP::HIS3MX6*	This study
RPY4961	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-7-GFP::HIS3MX6*	This study
RPY4962	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-8-GFP::HIS3MX6*	This study
RPY4980	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-9-GFP::HIS3MX6*	This study
RPY4981	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-10-GFP::HIS3MX6*	This study
RPY4982	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-11-GFP::HIS3MX6*	This study
RPY4983	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-9-GFP::HIS3MX6 pYES-97QmCh*	This study
RPY4984	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-10-GFP::HIS3MX6 pYES-97QmCh*	This study
RPY4985	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 SCP160KH1-11-GFP::HIS3MX6 pYES-97QmCh*	This study
RPY4990	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97QmCh::URA3 pUG34-GFP*	This study
RPY4993	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 NAB3-GFP::HIS3MX6*	This study
RPY4994	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-97QmCh::URA3 pUG34-GFP*	This study
RPY5000	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 NAB3-GFP::HIS3MX6*	This study

RPY5072	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 NAB3-3xmyc::HIS3MX6*	This study
RPY5131	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-(-)ATG-97Q(CAG)mCh::URA3*	This study
RPY5132	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-(-)ATG-97Q(CAG)mCh::URA3*	This study
RPY5180	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::KANMX6 NAB3-3xmyc::HIS3MX6*	This study
RPY5221	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-97QmCh::URA3 pRS313-SCP160*	This study
RPY5222	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-97QmCh::URA3 pRS313-SCP160*	This study
RPY5223	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97QmCh::URA3 pRS313-SCP160*	This study
RPY5224	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97QmCh::URA3 pRS313*	This study
RPY5226	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 ltn1::HIS3MX6 pYES-97QmCh::URA3*	This study
RPY5229	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 ltn1::HIS3MX6 pYES-97QmCh::URA3*	This study
RPY5271	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-92Q(CAA)mCh::URA3*	This study
RPY5272	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-92Q(CAA)mCh::URA3*	This study

5.4 – Experimental procedures

Experimental procedures for yeast work, induction of polyQ reporters, serial dilution assays, filter-trap binding, Scp160p depletion by Tet-off system, and mass spectrometry are described in the Research Article.

5.4.1 – Ploidy analysis by propidium iodide and flow cytometry

To collect cells for flow cytometry, yeast cultures were inoculated in YEPD to an OD_{600} of approximately 0.20 and grown at 30°C for 24 hours with aeration. After 24 hours of growth, 2OD units of cells were collected, washed twice with 50mM Tris-HCl pH 8.0, fixed for 1hr at room temperature in 1mL 50mM Tris-HCl pH 8.0 containing 70% ethanol, and stored at 4°C for up to one week before subsequent steps. 100uL of the fixed cells were pelleted by centrifugation for 10min at 17,115xg and washed twice with 1mL 50mM Tris-HCl pH 8.0. Cell pellets were resuspended in 0.5mL 50mM Tris-HCl pH 8.0 containing 20µg/mL heat-treated RNase A at 50°C for 2hr with gentle shaking. 20µL of Proteinase K (20mg/mL, Thermo Fischer Scientific #EO0491) at 50°C for 1hr. Cells were pelleted by centrifugation at 17,115xg for 10min, washed 3x with 1mL FACS buffer (200mM Tris-HCl pH 7.5, 211mM NaCl, 78mM $MgCl_2$), and stored at 4°C in FACS buffer. 500µL aliquots of the cells were pelleted by centrifugation at 17,115xg for 10min and resuspended in FACS-PI buffer (180mM Tris-HCl pH 7.5, 190mM NaCl, 70mM $MgCl_2$, 100µg/mL propidium iodide) and stained for 1-2hrs at room temperature on a wheel in the dark. Samples were vortexed before acquisitions were made on the flow cytometer.

Acquisitions were made on a Beckmann Coulter CytoFLEX S flow cytometer equipped with 488nm, 638nm, 405nm, and 561nm lasers, and BP 510/20 GFP, BP 515/20 eGFP, BP 595/20 DsRed, BP 585/15 dTomato, and BP 610/20 mCherry filters. Propidium iodide signal was measured using the ECD channel (638nm laser with BP 610/20 mCherry filter) with a gain of 163. Signal was recorded for 30,000 events or 10min at a flow rate of 30µL/min. Cell debris was excluded from analysis by gating from an FSC vs SSC density plot (P1). Cell clumps and doublets were further excluded from analysis by gating the P1 population using an ECD-Width (x-axis) vs ECD-Area (y-axis) density plot (P2).

5.4.2 – Protein work and co-immunoprecipitation

Protein lysates were made from logarithmically growing cells by glass bead lysis. The appropriate OD_{600} unit of cells were collected by centrifugation at 1,258xg for 5min at room temperature. The cell pellets were washed with sterile water and flash frozen in liquid nitrogen. To lyse, cell pellets were resuspended in 15µL/OD_{600} lysis buffer (25mM Tris-HCl pH 7.5, 50mM KCl, 10mM $MgCl_2$, 1mM EDTA, 5% glycerol, 0.5% triton X-100, 1x protease inhibitors (Roche cOmplete, EDTA-free tablets)) and

lysed with glass beads by vortexing for 6x 2min cycles with 1min rest on ice in between. Glass beads and cell debris were separated from the lysate by centrifugation at 1,200xg for 2min at 4°C. Protein concentrations were determined by Bradford assay and lysates were stored at -20°C until use.

Methods for western blotting and filter-trap binding were performed according to standard methods and have been described in the Research Article. For immunoprecipitation experiments, 200-240µg of lysates were incubated with GFP-Trap magnetic agarose beads (ChromoTek) according to manufacturer's protocol. Briefly, 12-25µL of bead slurry was equilibrated by addition to 500µL dilution/wash buffer (10mM Tris-HCl pH 7.5, 150mM NaCl, 0.5mM EDTA). The beads were magnetically separated and washed twice more with dilution/wash buffer. Lysates diluted to 500µL in dilution/wash buffer were added to equilibrated beads. Binding proceeded for at least 1 hour at 4°C with end-over-end mixing. After binding, beads were magnetically separated and washed twice. Bound proteins were eluted from the beads by boiling at 95°C for 5min in 50µL of 1x Laemmli buffer.

5.4.3 – RNA extraction and RT-qPCR

To extract total RNA, cell pellets were re-suspended in lysis buffer (0.3M NaCl, 10mM Tris-HCl pH 7.5, 1mM EDTA, 0.2% SDS) and lysed by glass beads in the presence of phenol/chloroform. Lysis was carried out at 4°C by vortexing for 10min. RNA was precipitated from the aqueous phase with the addition of RNase-free ethanol. Pelleted RNAs were re-suspended in DEPC treated water and stored at -20°C.

For RT-qPCR, 1µg of RNA was treated with RQ1 DNase (Promega) according to manufacturer's protocol. DNase treated RNAs were subsequently used for cDNA synthesis using the High-Capacity cDNA Reverse Transcription Kit (Thermofisher). qPCR was performed in a OneStepPlus Real-Time PCR system with Fast SYBR Green Master Mix (Applied Biosystems). The polyQ reporter mRNAs were detected using primers targeting the mCherry-encoding nucleotide sequence. Relative RNA abundance was calculated using the $\Delta\Delta C_t$ method with 18S or *ACT1* as endogenous controls.

5.4.4 – Mild cycloheximide treatment

To perturb translation, cultures were treated with a low dose of cycloheximide. The concentration of cycloheximide used (30ng/mL or 15ng/mL) was in the range reported to induce ribosome collision (Simms et al., 2017) and additionally determined by testing (data not shown). For the 97QmCh experiments, saturated overnight cultures in selective media + 2% glucose were inoculated into selective media + 1% raffinose and allowed to recover at 30°C for 1 hour with aeration. 97QmCh expression was induced by addition of galactose to a final concentration of 2%. After 1 hour of induction, cycloheximide was added to the cultures to a final concentration of 30ng/mL. Cells were collected at 3 hours after induction (= 2 hours after cycloheximide addition) and 7 hours after induction (= 6 hours after cycloheximide addition). Cells were collected as described above.

6. REFERENCES

Akopian, D., Shen, K., Zhang, X., and Shan, S. (2013). Signal Recognition Particle: An Essential Protein-Targeting Machine. Annu. Rev. Biochem. *82*, 693–721, 10.1146/annurev-biochem-072711-164732.

Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins. Cell *137*, 146–158, 10.1016/j.cell.2009.02.044.

Alexandrov, A.I., and Ter-Avanesyan, M.D. (2013). Could yeast prion domains originate from polyQ/N tracts? Prion *7*, 209–214, 10.4161/pri.24628.

Aronov, S., Gelin-Licht, R., Zipor, G., Haim, L., Safran, E., and Gerst, J.E. (2007). mRNAs encoding polarity and exocytosis factors are cotransported with the cortical endoplasmic reticulum to the incipient bud in Saccharomyces cerevisiae. Mol. Cell. Biol. *27*, 3441–3455, 10.1128/MCB.01643-06.

Aslam, K., and Hazbun, T.R. (2016). Hsp31, a member of the DJ-1 superfamily, is a multitasking stress responder with chaperone activity. Prion *10*, 103–111, 10.1080/19336896.2016.1141858.

Atanesyan, L., Günther, V., Dichtl, B., Georgiev, O., and Schaffner, W. (2012). Polyglutamine tracts as modulators of transcriptional activation from yeast to mammals. Biol. Chem. *393*, 63–70, 10.1515/BC-2011-252.

Athey, J., Alexaki, A., Osipova, E., Rostovtsev, A., Santana-Quintero, L. V, Katneni, U., Simonyan, V., and Kimchi-Sarfaty, C. (2017). A new and updated resource for codon usage tables. BMC Bioinformatics *18*, 391, 10.1186/s12859-017-1793-7.

Bartel, D.P. (2018). Metazoan MicroRNAs. Cell *173*, 20–51, 10.1016/j.cell.2018.03.006.

Bäuerlein, F.J.B., Saha, I., Mishra, A., Kalemanov, M., Martínez-Sánchez, A., Klein, R., Dudanova, I., Hipp, M.S., Hartl, F.U., Baumeister, W., et al. (2017). In Situ Architecture and Cellular Interactions of PolyQ Inclusions. Cell *171*, 179–187.e10, 10.1016/j.cell.2017.08.009.

Baum, S., Bittins, M., Frey, S., and Seedorf, M. (2004). Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes. Biochem. J. *380*, 823–830, 10.1042/BJ20031962.

Bazzini, A.A., Del Viso, F., Moreno-Mateos, M.A., Johnstone, T.G., Vejnar, C.E., Qin, Y., Yao, J., Khokha, M.K., and Giraldez, A.J. (2016). Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. EMBO J. *35*, 2087–2103, 10.15252/embj.201694699.

Becker, A.H., Oh, E., Weissman, J.S., Kramer, G., and Bukau, B. (2013). Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. Nat. Protoc. *8*, 2212–2239, 10.1038/nprot.2013.133.

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The structure of the eukaryotic ribosome at 3.0 Å resolution. Science *334*, 1524–1529, 10.1126/science.1212642.

Bennetzen, J.L., and Hall, B.D. (1982). Codon selection in yeast. J. Biol. Chem. *257*, 3026–3031.

Bergeron-Sandoval, L.-P., Safaee, N., and Michnick, S.W. (2016). Mechanisms and Consequences of Macromolecular Phase Separation. Cell *165*, 1067–1079, 10.1016/j.cell.2016.05.026.

Berhani, O., Nachmani, D., Yamin, R., Schmiedel, D., Bar-On, Y., and Mandelboim, O. (2017). Vigilin Regulates the Expression of the Stress-Induced Ligand MICB by Interacting with Its 5' Untranslated Region. J. Immunol. *198*, 3662–3670, 10.4049/jimmunol.1601589.

Bian, Z., Ni, Y., Xu, J.-R., and Liu, H. (2018). A-to-I mRNA editing in fungi: occurrence, function, and evolution. Cell. Mol. Life Sci. 1–12, 10.1007/s00018-018-2936-3.

Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. (1987). 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. *154*, 164–175, 10.1016/0076-6879(87)54076-9.

de Boer, P., Hoogenboom, J.P., and Giepmans, B.N. (2015). Correlated light and electron microscopy: ultrastructure lights up! Nat. Methods *12*, 503–513, 10.1038/nmeth.3400.

Boke, E., Ruer, M., Wühr, M., Coughlin, M., Lemaitre, R., Gygi, S.P., Alberti, S., Drechsel, D., Hyman, A.A., and Mitchison, T.J. (2016). Amyloid-like Self-Assembly of a Cellular Compartment. Cell *166*, 637–650, 10.1016/j.cell.2016.06.051.

Brandman, O., and Hegde, R.S. (2016). Ribosome-associated protein quality control. Nat. Struct. Mol. Biol. *23*, 7–15, 10.1038/nsmb.3147.

Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.-W., Zhou, S., King, D., Shen, P.S., Weibezahn, J., et al. (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell *151*, 1042–1054, 10.1016/j.cell.2012.10.044.

Brangwynne, C.P., Tompa, P., and Pappu, R. V (2015). Polymer physics of intracellular phase transitions. Nat. Phys. *11*, 899–904, 10.1038/nphys3532.

Buhr, F., Jha, S., Thommen, M., Mittelstaet, J., Kutz, F., Schwalbe, H., Rodnina, M. V, and Komar, A.A. (2016). Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations. Mol. Cell *61*, 341–351, 10.1016/j.molcel.2016.01.008.

Cannarozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G., and Barral, Y. (2010). A role for codon order in translation dynamics. Cell *141*, 355–367, 10.1016/j.cell.2010.02.036.

Cheng, M.H.K., and Jansen, R.-P. (2017). A jack of all trades: the RNA-binding protein

vigilin. Wiley Interdiscip. Rev. RNA *8*, e1448, 10.1002/wrna.1448.

Cheng, M.H.K., Hoffmann, P.C., Franz-Wachtel, M., Sparn, C., Seng, C., Maček, B., and Jansen, R.-P. (2018). The RNA-Binding Protein Scp160p Facilitates Aggregation of Many Endogenous Q/N-Rich Proteins. Cell Rep. *24*, 20–26, 10.1016/j.celrep.2018.06.015.

Choe, Y.-J., Park, S.-H., Hassemer, T., Körner, R., Vincenz-Donnelly, L., Hayer-Hartl, M., and Hartl, F.U. (2016). Failure of RQC machinery causes protein aggregation and proteotoxic stress. Nature *531*, 191–195, 10.1038/nature16973.

Ciechanover, A., and Kwon, Y.T. (2017). Protein Quality Control by Molecular Chaperones in Neurodegeneration. Front. Neurosci. *11*, 185, 10.3389/fnins.2017.00185.

Cléry, A., and Allain, F.H.-T. (2012). From Structure To Function of RNA Binding Domains. In Madame Curie Bioscience Database [Internet], (Austin (TX): Landes Bioscience), pp. 137–158.

Cortés, A., Huertas, D., Fanti, L., Pimpinelli, S., Marsellach, F.X., Piña, B., and Azorín, F. (1999). DDP1, a single-stranded nucleic acid-binding protein of Drosophila, associates with pericentric heterochromatin and is functionally homologous to the yeast Scp160p, which is involved in the control of cell ploidy. EMBO J. *18*, 3820–3833, 10.1093/emboj/18.13.3820.

Crick, F.H.C. (1966). Codon--anticodon pairing: the wobble hypothesis. J. Mol. Biol. *19*, 548–555, 10.1016/S0022-2836(66)80022-0.

Cunningham, K.S., Dodson, R.E., Nagel, M.A., Shapiro, D.J., and Schoenberg, D.R. (2000). Vigilin binding selectively inhibits cleavage of the vitellogenin mRNA 3'-untranslated region by the mRNA endonuclease polysomal ribonuclease 1. Proc. Natl. Acad. Sci. U. S. A. *97*, 12498–12502, 10.1073/pnas.220425497.

Darnell, A.M., Subramaniam, A.R., and O'Shea, E.K. (2018). Translational Control through Differential Ribosome Pausing during Amino Acid Limitation in Mammalian Cells. Mol. Cell *71*, 229–243.e11, 10.1016/j.molcel.2018.06.041.

Decker, C.J., and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb. Perspect. Biol. *4*, a012286, 10.1101/cshperspect.a012286.

Decker, C.J., Teixeira, D., and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. J. Cell Biol. *179*, 437–449, 10.1083/jcb.200704147.

Deng, Y., Singer, R.H., and Gu, W. (2008). Translation of ASH1 mRNA is repressed by Puf6p-Fun12p/eIF5B interaction and released by CK2 phosphorylation. Genes Dev. *22*, 1037–1050, 10.1101/gad.1611308.

Dittmar, K.A., Sørensen, M.A., Elf, J., Ehrenberg, M., and Pan, T. (2005). Selective charging of tRNA isoacceptors induced by amino-acid starvation. EMBO Rep. *6*, 151–157, 10.1038/sj.embor.7400341.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561–564, 10.1038/nature04530.

Duennwald, M.L., and Lindquist, S. (2008). Impaired ERAD and ER stress are early and specific events in polyglutamine toxicity. Genes Dev. *22*, 3308–3319, 10.1101/gad.1673408.

Duennwald, M.L., Jagadish, S., Muchowski, P.J., and Lindquist, S. (2006). Flanking sequences profoundly alter polyglutamine toxicity in yeast. Proc. Natl. Acad. Sci. U. S. A. *103*, 11045–11050, 10.1073/pnas.0604547103.

Duennwald, M.L., Echeverria, A., and Shorter, J. (2012). Small heat shock proteins potentiate amyloid dissolution by protein disaggregases from yeast and humans. PLoS Biol. *10*, e1001346, 10.1371/journal.pbio.1001346.

Duret, L. (2000). tRNA gene number and codon usage in the C. elegans genome are co-adapted for optimal translation of highly expressed genes. Trends Genet. *16*, 287–289, 10.1016/S0168-9525(00)02041-2.

Fan, H., Ho, L., Chi, C., Chen, S., Peng, G., Chan, T., Lin, S., and Harn, H. (2014). Polyglutamine (PolyQ) diseases: genetics to treatments. Cell Transplant. *23*, 441–458, 10.3727/096368914X678454.

Fiumara, F., Fioriti, L., Kandel, E.R., and Hendrickson, W.A. (2010). Essential role of coiled coils for aggregation and activity of Q/N-rich prions and PolyQ proteins. Cell *143*, 1121–1135, 10.1016/j.cell.2010.11.042.

Fluitt, A., Pienaar, E., and Viljoen, H. (2007). Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis. Comput. Biol. Chem. *31*, 335–346, 10.1016/j.compbiolchem.2007.07.003.

Franzmann, T.M., Jahnel, M., Pozniakovsky, A., Mahamid, J., Holehouse, A.S., Nüske, E., Richter, D., Baumeister, W., Grill, S.W., Pappu, R. V, et al. (2018). Phase separation of a yeast prion protein promotes cellular fitness. Science *359*, 10.1126/science.aao5654.

Frey, S., Pool, M., and Seedorf, M. (2001). Scp160p, an RNA-binding, polysome-associated protein, localizes to the endoplasmic reticulum of Saccharomyces cerevisiae in a microtubule-dependent manner. J. Biol. Chem. *276*, 15905–15912, 10.1074/jbc.M009430200.

Fu, J., Murphy, K.A., Zhou, M., Li, Y.H., Lam, V.H., Tabuloc, C.A., Chiu, J.C., and Liu, Y. (2016). Codon usage affects the structure and function of the Drosophila circadian clock protein PERIOD. Genes Dev. *30*, 1761–1775, 10.1101/gad.281030.116.

Fundakowski, J., Hermesh, O., and Jansen, R.-P. (2012). Localization of a subset of yeast mRNAs depends on inheritance of endoplasmic reticulum. Traffic *13*, 1642–1652, 10.1111/tra.12011.

Gavin, A.-C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631–636, 10.1038/nature04532.

Gelin-Licht, R., Paliwal, S., Conlon, P., Levchenko, A., and Gerst, J.E. (2012). Scp160-Dependent mRNA Trafficking Mediates Pheromone Gradient Sensing and Chemotropism in Yeast. Cell Rep. *1*, 483–494, 10.1016/j.celrep.2012.03.004.

Gemayel, R., Chavali, S., Pougach, K., Legendre, M., Zhu, B., Boeynaems, S., van der Zande, E., Gevaert, K., Rousseau, F., Schymkowitz, J., et al. (2015). Variable Glutamine-Rich Repeats Modulate Transcription Factor Activity. Mol. Cell *59*, 615–627, 10.1016/j.molcel.2015.07.003.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. Mol. Biol. Cell *15*, 5383–5398, 10.1091/mbc.E04-08-0715.

Gingold, H., Tehler, D., Christoffersen, N.R., Nielsen, M.M., Asmar, F., Kooistra, S.M., Christophersen, N.S., Christensen, L.L., Borre, M., Sørensen, K.D., et al. (2014). A dual program for translation regulation in cellular proliferation and differentiation. Cell *158*, 1281–1292, 10.1016/j.cell.2014.08.011.

Giorgini, F., Guidetti, P., Nguyen, Q., Bennett, S.C., and Muchowski, P.J. (2005). A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for Huntington disease. Nat. Genet. *37*, 526–531, 10.1038/ng1542.

Glover, J.R., Kowal, A.S., Schirmer, E.C., Patino, M.M., Liu, J.-J., and Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of S. cerevisiae. Cell *89*, 811–819, 10.1016/S0092-8674(00)80264-0.

Gouy, M., and Gautier, C. (1982). Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. *10*, 7055–7074, 10.1093/nar/10.22.7055.

Graham, D.L., and Oram, J.F. (1987). Identification and characterization of a high density lipoprotein-binding protein in cell membranes by ligand blotting. J. Biol. Chem. *262*, 7439–7442.

Grima, J.C., Daigle, J.G., Arbez, N., Cunningham, K.C., Zhang, K., Ochaba, J., Geater, C., Morozko, E., Stocksdale, J., Glatzer, J.C., et al. (2017). Mutant Huntingtin Disrupts the Nuclear Pore Complex. Neuron *94*, 93–107.e6, 10.1016/j.neuron.2017.03.023.

Harbi, D., and Harrison, P.M. (2014). Interaction networks of prion, prionogenic and prion-like proteins in budding yeast, and their role in gene regulation. PLoS One *9*, e100615, 10.1371/journal.pone.0100615.

Harigaya, Y., and Parker, R. (2017). The link between adjacent codon pairs and mRNA stability. BMC Genomics *18*, 364, 10.1186/s12864-017-3749-8.

Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. Nat. Rev. Mol. Cell Biol. 10.1038/nrm.2017.130.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harb. Perspect. Biol. *4*, 1–25, 10.1101/cshperspect.a011544.

Hirschmann, W.D., Westendorf, H., Mayer, A., Cannarozzi, G., Cramer, P., and Jansen, R.-P. (2014). Scp160p is required for translational efficiency of codon-optimized mRNAs in yeast. Nucleic Acids Res. *42*, 4043–4055, 10.1093/nar/gkt1392.

Ho, C.S., Khadka, N.K., She, F., Cai, J., and Pan, J. (2016). Polyglutamine aggregates impair lipid membrane integrity and enhance lipid membrane rigidity. Biochim. Biophys. Acta *1858*, 661–670, 10.1016/j.bbamem.2016.01.016.

Hofer, S., Kainz, K., Zimmermann, A., Bauer, M.A., Pendl, T., Poglitsch, M., Madeo, F., and Carmona-Gutierrez, D. (2018). Studying Huntington's Disease in Yeast: From Mechanisms to Pharmacological Approaches. Front. Mol. Neurosci. *11*, 318, 10.3389/fnmol.2018.00318.

Hoffner, G., and Djian, P. (2014). Monomeric, Oligomeric and Polymeric Proteins in Huntington Disease and Other Diseases of Polyglutamine Expansion. Brain Sci. *4*, 91–122, 10.3390/brainsci4010091.

Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D., and Brown, P.O. (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol. *6*, e255, 10.1371/journal.pbio.0060255.

Huertas, D., Cortés, A., Casanova, J., and Azorín, F. (2004). Drosophila DDP1, a multi-KH-domain protein, contributes to centromeric silencing and chromosome segregation. Curr. Biol. *14*, 1611–1620, 10.1016/j.cub.2004.09.024.

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. *2*, 13–34, 10.1093/oxfordjournals.molbev.a040335.

Ikeuchi, K., and Inada, T. (2016). Ribosome-associated Asc1/RACK1 is required for endonucleolytic cleavage induced by stalled ribosome at the 3' end of nonstop mRNA. Sci. Rep. *6*, 28234, 10.1038/srep28234.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218–223, 10.1126/science.1168978.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat. Protoc. *7*, 1534–1550, 10.1038/nprot.2012.086.

Jackson, C.L., and Képès, F. (1994). BFR1, a multicopy suppressor of brefeldin A-induced lethality, is implicated in secretion and nuclear segregation in Saccharomyces cerevisiae. Genetics *137*, 423–437.

Jackson, R.J., Hellen, C.U.T., and Pestova, T. V (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. *11*, 113–127, 10.1038/nrm2838.

Jacobson, G.N., and Clark, P.L. (2016). Quality over quantity: optimizing co-translational protein folding with non-'optimal' synonymous codons. Curr. Opin. Struct.

Biol. *38*, 102–110, 10.1016/j.sbi.2016.06.002.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., et al. (2004). A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947–962, 10.1002/yea.1142.

Jansen, R.-P., and Niessing, D. (2012). Assembly of mRNA-Protein Complexes for Directional mRNA Transport in Eukaryotes - An Overview. Curr. Protein Pept. Sci. *13*, 284–293, 10.2174/138920312801619493.

Jansen, R.-P., Dowzer, C., Michaelis, C., Galova, M., and Nasmyth, K. (1996). Mother cell-specific HO expression in budding yeast depends on the unconventional myosin Myo4p and other cytoplasmic proteins. Cell *84*, 687–697, 10.1016/S0092-8674(00)81047-8.

Kaganovich, D., Kopito, R., and Frydman, J. (2008). Misfolded proteins partition between two distinct quality control compartments. Nature *454*, 1088–1095, 10.1038/nature07195.

Kaiser, C.J.O., Grötzinger, S.W., Eckl, J.M., Papsdorf, K., Jordan, S., and Richter, K. (2013). A network of genes connects polyglutamine toxicity to ploidy control in yeast. Nat. Commun. *4*, 1571, 10.1038/ncomms2575.

Kayatekin, C., Matlack, K.E., Hesse, W.R., Guan, Y., Chakrabortee, S., Russ, J., Wanker, E.E., Shah, J. V, and Lindquist, S. (2014). Prion-like proteins sequester and suppress the toxicity of huntingtin exon 1. Proc. Natl. Acad. Sci. U. S. A. 1–6, 10.1073/pnas.1412504111.

Khan, M.R., Li, L., Pérez-Sánchez, C., Saraf, A., Florens, L., Slaughter, B.D., Unruh, J.R., and Si, K. (2015). Amyloidogenic Oligomerization Transforms Drosophila Orb2 from a Translation Repressor to an Activator. Cell *163*, 1468–1483, 10.1016/j.cell.2015.11.020.

Kim, H., and Strange, K. (2013). Changes in translation rate modulate stress-induced damage of diverse proteins. Am. J. Physiol. Cell Physiol. *305*, C1257-64, 10.1152/ajpcell.00176.2013.

Kim, Y.E., Hosp, F., Mann, M., Hayer-hartl, M., and Hartl, F.U. (2016). Soluble Oligomers of PolyQ-Expanded Huntingtin Target a Multiplicity of Key Cellular Factors. Mol. Cell 1–14, 10.1016/j.molcel.2016.07.022.

Kimata, Y., Ishiwata-Kimata, Y., Yamada, S., and Kohno, K. (2006). Yeast unfolded protein response pathway regulates expression of genes for anti-oxidative stress and for cell surface proteins. Genes Cells *11*, 59–69, 10.1111/j.1365-2443.2005.00921.x.

Kimchi-Sarfaty, C., Oh, J.M., Kim, I.-W., Sauna, Z.E., Calcagno, A.M., Ambudkar, S. V, and Gottesman, M.M. (2007). A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science *315*, 525–528, 10.1126/science.1135308.

King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res. *1462*, 61–

80, 10.1016/j.brainres.2012.01.016.

Kitamura, E., Tanaka, K., Kitamura, Y., and Tanaka, T.U. (2007). Kinetochore-microtubule interaction during S phase in Saccharomyces cerevisiae. Genes Dev. *21*, 3319–3330, 10.1101/gad.449407.

Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K., and Schiebel, E. (1999). Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. Yeast *15*, 963–972, 10.1002/(SICI)1097-0061(199907)15:10B<963::AID-YEA399>3.0.CO;2-W.

Kramer, K., Sachsenberg, T., Beckmann, B.M., Qamar, S., Boon, K.-L., Hentze, M.W., Kohlbacher, O., and Urlaub, H. (2014). Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. Nat. Methods *11*, 1064–1070, 10.1038/nmeth.3092.

Kruse, C., Willkomm, D.K., Grünweller, A., Vollbrandt, T., Sommer, S., Busch, S., Pfeiffer, T., Brinkmann, J., Hartmann, R.K., and Müller, P.K. (2000). Export and transport of tRNA are coupled to a multi-protein complex. Biochem. J. *346 Pt 1*, 107–115, 10.1042/0264-6021:3460107.

Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L., and McKnight, S.L. (2013). Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell *155*, 1049–1060, 10.1016/j.cell.2013.10.033.

Lang, B.D., and Fridovich-Keil, J.L. (2000). Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. Nucleic Acids Res. *28*, 1576–1584, gkd275 [pii].

Lang, B.D., Li, A.-M., Black-Brewster, H.D., and Fridovich-Keil, J.L. (2001). The brefeldin A resistance protein Bfr1p is a component of polyribosome-associated mRNP complexes in yeast. Nucleic Acids Res. *29*, 2567–2574, 10.1093/nar/29.12.2567.

Langdon, E.M., Qiu, Y., Ghanbari Niaki, A., McLaughlin, G.A., Weidmann, C.A., Gerbich, T.M., Smith, J.A., Crutchley, J.M., Termini, C.M., Weeks, K.M., et al. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. Science *360*, 922–927, 10.1126/science.aar7432.

Lapointe, C.P., Wilinski, D., Saunders, H.A., and Wickens, M. (2015). Protein-RNA networks revealed through covalent RNA marks. Nat. Methods *12*, 1163–1170, 10.1038/nmeth.3651.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature *547*, 236–240, 10.1038/nature22822.

Lasko, P. (2012). mRNA localization and translational control in Drosophila oogenesis. Cold Spring Harb. Perspect. Biol. *4*, 1–15, 10.1101/cshperspect.a012294.

Lee, C., Zhang, H., Baker, A.E., Occhipinti, P., Borsuk, M.E., and Gladfelter, A.S. (2013). Protein aggregation behavior regulates cyclin transcript localization and cell-cycle control. Dev. Cell *25*, 572–584, 10.1016/j.devcel.2013.05.007.

Lee, C., Occhipinti, P., and Gladfelter, A.S. (2015). PolyQ-dependent RNA-protein assemblies control symmetry breaking. J. Cell Biol. *208*, 533–544, 10.1083/jcb.201407105.

Leitman, J., Hartl, F.U., and Lederkremer, G.Z. (2013). Soluble forms of polyQ-expanded huntingtin rather than large aggregates cause endoplasmic reticulum stress. Nat. Commun. *4*, 2753, 10.1038/ncomms3753.

Li, A.-M., Watson, A., and Fridovich-Keil, J.L. (2003). Scp160p associates with specific mRNAs in yeast. Nucleic Acids Res. *31*, 1830–1837, 10.1093/nar/gkg284.

Li, A.-M., Vargas, C.A., Brykailo, M.A., Openo, K.K., Corbett, A.H., and Fridovich-Keil, J.L. (2004). Both KH and non-KH domain sequences are required for polyribosome association of Scp160p in yeast. Nucleic Acids Res. *32*, 4768–4775, 10.1093/nar/gkh812.

Li, L., McGinnis, J., and Si, K. (2018). Translational Control by Prion-like Proteins. Trends Cell Biol. *28*, 494–505, 10.1016/j.tcb.2018.02.002.

Licht, K., and Jantsch, M.F. (2016). Rapid and dynamic transcriptome regulation by RNA editing and RNA modifications. J. Cell Biol. *213*, 15–22, 10.1083/jcb.201511041.

Liebman, S.W., and Chernoff, Y.O. (2012). Prions in yeast. Genetics *191*, 1041–1072, 10.1534/genetics.111.137760.

Lin, Y., Protter, D.S., Rosen, M.K., and Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol. Cell *60*, 208–219, 10.1016/j.molcel.2015.08.018.

Liu, H., Ma, C.-P., Chen, Y.-T., Schuyler, S.C., Chang, K.-P., and Tan, B.C.-M. (2014a). Functional Impact of RNA editing and ADARs on regulation of gene expression: perspectives from deep sequencing studies. Cell Biosci. *4*, 44, 10.1186/2045-3701-4-44.

Liu, Q., Yang, B., Xie, X., Wei, L., Liu, W., Yang, W., Ge, Y., Zhu, Q., Zhang, J., Jiang, L., et al. (2014b). Vigilin interacts with CCCTC-binding factor (CTCF) and is involved in CTCF-dependent regulation of the imprinted genes Igf2 and H19. FEBS J. *281*, 2713–2725, 10.1111/febs.12816.

Longtine, M.S., McKenzie, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953–961, 10.1002/(SICI)1097-0061(199807)14:10<953::AID-YEA293>3.0.CO;2-U.

Low, Y.S., Bircham, P.W., Maass, D.R., and Atkinson, P.H. (2014). Kinetochore genes are required to fully activate secretory pathway expansion in S. cerevisiae under induced ER stress. Mol. Biosyst. *10*, 1790–1802, 10.1039/c3mb70414a.

Loya, T.J., O'Rourke, T.W., Degtyareva, N., and Reines, D. (2013). A network of interdependent molecular interactions describes a higher order Nrd1-Nab3 complex involved in yeast transcription termination. J. Biol. Chem. *288*, 34158–34167, 10.1074/jbc.M113.516765.

Lu, S.H.-J., Jeon, A.H.W., Schmitt-Ulms, G., Qamar, S., Dodd, R., McDonald, B., Li, Y., Meadows, W., Cox, K., Bohm, C., et al. (2012). Vigilin interacts with signal peptide peptidase. Proteome Sci. *10*, 33, 10.1186/1477-5956-10-33.

Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat. Rev. Mol. Cell Biol. *8*, 479–490, 10.1038/nrm2178.

Macdonald, W.A. (2012). Epigenetic mechanisms of genomic imprinting: common themes in the regulation of imprinted regions in mammals, plants, and insects. Genet. Res. Int. *2012*, 585024, 10.1155/2012/585024.

Malinovska, L., Palm, S., Gibson, K., Verbavatz, J.-M., and Alberti, S. (2015). Dictyostelium discoideum has a highly Q/N-rich proteome and shows an unusual resilience to protein aggregation. Proc. Natl. Acad. Sci. U. S. A. *112*, E2620-9, 10.1073/pnas.1504459112.

Mar Albà, M., Santibáñez-Koref, M.F., and Hancock, J.M. (1999). Amino acid reiterations in yeast are overrepresented in particular classes of proteins and show evidence of a slippage-like mutational process. J. Mol. Evol. *49*, 789–797, 10.1007/PL00006601.

Marsellach, F.-X., Huertas, D., and Azorín, F. (2006). The multi-KH domain protein of Saccharomyces cerevisiae Scp160p contributes to the regulation of telomeric silencing. J. Biol. Chem. *281*, 18227–18235, 10.1074/jbc.M601671200.

Matsuda, R., Ikeuchi, K., Nomura, S., and Inada, T. (2014). Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. Genes to Cells *19*, 1–12, 10.1111/gtc.12106.

McKeehan, W., and Hardesty, B. (1969). The mechanism of cycloheximide inhibition of protein synthesis in rabbit reticulocytes. Biochem. Biophys. Res. Commun. *36*, 625–630, 10.1016/0006-291X(69)90351-9.

McKnight, G.L., Reasoner, J., Gilbert, T., Sundquist, K.O., Hokland, B., McKernan, P.A., Champagne, J., Johnson, C.J., Bailey, M.C., Holly, R., et al. (1992). Cloning and expression of a cellular high density lipoprotein-binding protein that is up-regulated by cholesterol loading of cells. J. Biol. Chem. *267*, 12131–12141.

Meriin, A.B., Zhang, X., Miliaras, N.B., Kazantsev, A., Chernoff, Y.O., McCaffery, J.M., Wendland, B., and Sherman, M.Y. (2003). Aggregation of expanded polyglutamine domain in yeast leads to defects in endocytosis. Mol. Cell. Biol. *23*, 7554–7565, 10.1128/MCB.23.21.7554-7565.2003.

Meriin, A.B., Mense, M., Colbert, J.D., Liang, F., Bihler, H., Zaarur, N., Rock, K.L., and Sherman, M.Y. (2012). A novel approach to recovery of function of mutant proteins by slowing down translation. J. Biol. Chem. *287*, 34264–34272, 10.1074/jbc.M112.397307.

de Mezer, M., Wojciechowska, M., Napierala, M., Sobczak, K., and Krzyzosiak, W.J. (2011). Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference. Nucleic Acids Res. *39*, 3852–3863, 10.1093/nar/gkq1323.

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., and Thomas, P.D. (2017). PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. *45*, D183–D189, 10.1093/nar/gkw1138.

Michelitsch, M.D., and Weissman, J.S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. Proc. Natl. Acad. Sci. U. S. A. *97*, 11910–11915, 10.1073/pnas.97.22.11910.

Miller, S.B.M., Mogk, A., and Bukau, B. (2015). Spatially organized aggregation of misfolded proteins as cellular stress defense strategy. J. Mol. Biol. *427*, 1564–1574, 10.1016/j.jmb.2015.02.006.

Mishima, Y., and Tomari, Y. (2016). Codon Usage and 3' UTR Length Determine Maternal mRNA Stability in Zebrafish. Mol. Cell *61*, 874–885, 10.1016/j.molcel.2016.02.027.

Mobin, M.B., Gerstberger, S., Teupser, D., Campana, B., Charisse, K., Heim, M.H., Manoharan, M., Tuschl, T., and Stoffel, M. (2016). The RNA-binding protein vigilin regulates VLDL secretion through modulation of Apob mRNA translation. Nat. Commun. *7*, 12848, 10.1038/ncomms12848.

Mogk, A., and Bukau, B. (2017). Role of sHsps in organizing cytosolic protein aggregation and disaggregation. Cell Stress Chaperones *22*, 493–502, 10.1007/s12192-017-0762-4.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. Cell *163*, 123–133, 10.1016/j.cell.2015.09.015.

Nalavade, R., Griesche, N., Ryan, D., Hildebrand, S., and Krauß, S. (2013). Mechanisms of RNA-induced toxicity in CAG repeat disorders. Cell Death Dis. *4*, e752, 10.1038/cddis.2013.276.

Negrutskii, B.S., and Deutscher, M.P. (1991). Channeling of aminoacyl-tRNA for protein synthesis in vivo. Proc. Natl. Acad. Sci. U. S. A. *88*, 4991–4995, 10.1073/pnas.88.11.4991.

Nelson, M.R., Luo, H., Vari, H.K., Cox, B.J., Simmonds, A.J., Krause, H.M., Lipshitz, H.D., and Smibert, C.A. (2007). A multiprotein complex that mediates translational enhancement in Drosophila. J. Biol. Chem. *282*, 34031–34038, 10.1074/jbc.M706363200.

Nicastro, G., Taylor, I.A., and Ramos, A. (2015). KH-RNA interactions: back in the groove. Curr. Opin. Struct. Biol. *30C*, 63–70, 10.1016/j.sbi.2015.01.002.

Nishimura, K., and Kanemaki, M.T. (2014). Rapid Depletion of Budding Yeast Proteins via the Fusion of an Auxin-Inducible Degron (AID). Curr. Protoc. Cell Biol. *64*, 20.9.1-16, 10.1002/0471143030.cb2009s64.

Nissley, D.A., and O'Brien, E.P. (2016). Altered Co-Translational Processing Plays a Role in Huntington's Pathogenesis-A Hypothesis. Front. Mol. Neurosci. *9*, 54, 10.3389/fnmol.2016.00054.

O'Rourke, T.W., and Reines, D. (2016). Determinants of Amyloid Formation for the Yeast Termination Factor Nab3. PLoS One *11*, e0150865, 10.1371/journal.pone.0150865.

Paquin, N., Ménade, M., Poirier, G., Donato, D., Drouet, E., and Chartrand, P. (2007). Local activation of yeast ASH1 mRNA translation through phosphorylation of Khd1p by the casein kinase Yck1p. Mol. Cell *26*, 795–809, 10.1016/j.molcel.2007.05.016.

Park, S.-H., Kukushkin, Y., Gupta, R., Chen, T., Konagai, A., Hipp, M.S., Hayer-Hartl, M., and Hartl, F.U. (2013). PolyQ proteins interfere with nuclear degradation of cytosolic proteins by sequestering the Sis1p chaperone. Cell *154*, 134–145, 10.1016/j.cell.2013.06.003.

Parker, R. (2012). RNA degradation in Saccharomyces cerevisae. Genetics *191*, 671–702, 10.1534/genetics.111.137265.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066–1077, 10.1016/j.cell.2015.07.047.

Patel, B.K., Gavin-Smyth, J., and Liebman, S.W. (2009). The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion. Nat. Cell Biol. *11*, 344–349, 10.1038/ncb1843.

Pechmann, S., Chartron, J.W., and Frydman, J. (2014). Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. Nat. Struct. Mol. Biol. *21*, 1100–1105, 10.1038/nsmb.2919.

Peskett, T.R., Rau, F., O'Driscoll, J., Patani, R., Lowe, A.R., and Saibil, H.R. (2018). A Liquid to Solid Phase Transition Underlying Pathological Huntingtin Exon1 Aggregation. Mol. Cell *70*, 588–601.e6, 10.1016/j.molcel.2018.04.007.

Pichon, X., Wilson, L.A., Stoneley, M., Bastide, A., King, H.A., Somers, J., and Willis, A.E. (2012). RNA Binding Protein/RNA Element Interactions and the Control of Translation. Curr. Protein Pept. Sci. *13*, 294–304, 10.2174/138920312801619475.

Porrua, O., and Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. Nat Rev Mol Cell Biol *16*, 190–202, 10.1038/nrm3943\rhttp://www.nature.com/nrm/journal/v16/n3/abs/nrm3943.html#supplementary-information.

Presnyak, V., Alhusaini, N., Chen, Y.-H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., et al. (2015). Codon optimality is a major

determinant of mRNA stability. Cell *160*, 1111–1124, 10.1016/j.cell.2015.02.029.

Purvis, I.J., Bettany, A.J.E., Santiago, T.C., Coggins, J.R., Duncan, K., Eason, R., and Brown, A.J.P. (1987). The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. A hypothesis. J. Mol. Biol. *193*, 413–417, 10.1016/0022-2836(87)90230-0.

Quax, T.E., Claassens, N.J., Söll, D., and van der Oost, J. (2015). Codon Bias as a Means to Fine-Tune Gene Expression. Mol. Cell *59*, 149–161, 10.1016/j.molcel.2015.05.035.

Rabouille, C., and Alberti, S. (2017). Cell adaptation upon stress: the emerging role of membrane-less compartments. Curr. Opin. Cell Biol. *47*, 34–42, 10.1016/j.ceb.2017.02.006.

Radhakrishnan, A., Chen, Y.-H., Martin, S., Alhusaini, N., Green, R., and Coller, J. (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell *167*, 122–132.e9, 10.1016/j.cell.2016.08.053.

Raulet, D.H., Gasser, S., Gowen, B.G., Deng, W., and Jung, H. (2013). Regulation of ligands for the NKG2D activating receptor. Annu. Rev. Immunol. *31*, 413–441, 10.1146/annurev-immunol-032712-095951.

Reijns, M.A., Alexander, R.D., Spiller, M.P., and Beggs, J.D. (2008). A role for Q/N-rich aggregation-prone regions in P-body localization. J. Cell Sci. *121*, 2463–2472, 10.1242/jcs.024976.

Rodnina, M. V (2016). The ribosome in action: Tuning of translational efficiency and protein folding. Protein Sci. *25*, 1390–1406, 10.1002/pro.2950.

Saad, S., Cereghetti, G., Feng, Y., Picotti, P., Peter, M., and Dechant, R. (2017). Reversible protein aggregation is a protective mechanism to ensure cell cycle restart after stress. Nat. Cell Biol. *19*, 1202–1213, 10.1038/ncb3600.

Saikia, M., Wang, X., Mao, Y., Wan, J., Pan, T., and Qian, S.-B. (2016). Codon optimality controls differential mRNA translation during amino acid starvation. RNA *22*, 1719–1727, 10.1261/rna.058180.116.

Sakahira, H., Breuer, P., Hayer-Hartl, M.K., and Hartl, F.U. (2002). Molecular chaperones as modulators of polyglutamine protein aggregation and toxicity. Proc. Natl. Acad. Sci. U. S. A. *99 Suppl 4*, 16412–16418, 10.1073/pnas.182426899.

Schaefer, M.H., Wanker, E.E., and Andrade-Navarro, M.A. (2012). Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks. Nucleic Acids Res. *40*, 4273–4287, 10.1093/nar/gks011.

Scherzinger, E., Sittler, A., Schweiger, K., Heiser, V., Lurz, R., Hasenbank, R., Bates, G., Lehrach, H., and Wanker, E. (1999). Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. Proc. Natl. Acad. Sci. U. S. A. *96*, 4604–4609, 10.1073/PNAS.96.8.4604.

Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green,

R., Shen, B., and Liu, J.O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. Nat. Chem. Biol. *6*, 209–217, 10.1038/nchembio.304.

Schreck, H. (2010). Translational control by the multi KH domain protein Scp160. Ludwig-Maximilians-Universität München.

Schulte, J., and Littleton, J.T. (2011). The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology. Curr. Trends Neurol. *5*, 65–78.

Semotok, J.L., Cooperstock, R.L., Pinder, B.D., Vari, H.K., Lipshitz, H.D., and Smibert, C.A. (2005). Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early Drosophila embryo. Curr. Biol. *15*, 284–294, 10.1016/j.cub.2005.01.048.

Serpionov, G. V, Alexandrov, A.I., and Ter-Avanesyan, M.D. (2017). Distinct mechanisms of mutant huntingtin toxicity in different yeast strains. FEMS Yeast Res. *17*, 1–6, 10.1093/femsyr/fow102.

Sezen, B., Seedorf, M., and Schiebel, E. (2009). The SESA network links duplication of the yeast centrosome with the protein translation machinery. Genes Dev. *23*, 1559–1570, 10.1101/gad.524209.

Sharp, P.M., and Li, W.H. (1987). The codon adaptation index - A measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. *15*, 1281–1295, 10.1093/nar/15.3.1281.

Shen, P.S., Park, J., Qin, Y., Li, X., Parsawar, K., Larson, M.H., Cox, J., Cheng, Y., Lambowitz, A.M., Weissman, J.S., et al. (2015). Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. Science (80-.). *347*, 75–78, 10.1126/science.1259724.

Shen, W.Y., Liu, Q.Y., Wei, L., Yu, X.Q., Li, R., Yang, W.L., Xie, X.Y., Liu, W.Q., Huang, Y., and Qin, Y. (2014). CTCF-mediated reduction of vigilin binding affects the binding of HP1?? to the satellite 2 locus. FEBS Lett. *588*, 1549–1555, 10.1016/j.febslet.2014.02.013.

Sherman, M.Y., and Qian, S.-B. (2013). Less is more: improving proteostasis by translation slow down. Trends Biochem. Sci. *38*, 585–591, 10.1016/j.tibs.2013.09.003.

Shiber, A., Döring, K., Friedrich, U., Klann, K., Merker, D., Zedan, M., Tippmann, F., Kramer, G., and Bukau, B. (2018). Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. Nature 10.1038/s41586-018-0462-y.

Shoemaker, C.J., and Green, R. (2012). Translation drives mRNA quality control. Nat. Struct. Mol. Biol. *19*, 594–601, 10.1038/nsmb.2301.

Simms, C.L., Yan, L.L., and Zaher, H.S. (2017). Ribosome Collision Is Critical for Quality Control during No-Go Decay. Mol. Cell *68*, 361–373.e5, 10.1016/j.molcel.2017.08.019.

Singer-Krüger, B., and Jansen, R.-P. (2014). Here, there, everywhere: mRNA

localization in budding yeast. RNA Biol. *11*, 37–41, 10.4161/rna.29945.

Smith, R.L., and Johnson, A.D. (2000). Turning genes off by Ssn6-Tup1: a conserved system of transcriptional repression in eukaryotes. Trends Biochem. Sci. *25*, 325–330, 10.1016/S0968-0004(00)01592-9.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell *136*, 731–745, 10.1016/j.cell.2009.01.042.

Spencer, P.S., Siller, E., Anderson, J.F., and Barral, J.M. (2012). Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. J. Mol. Biol. *422*, 328–335, 10.1016/j.jmb.2012.06.010.

Strom, A.R., Emelyanov, A. V, Mir, M., Fyodorov, D. V, Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. Nature *547*, 241–245, 10.1038/nature22989.

Subramaniam, A.R., Zid, B.M., and O'Shea, E.K. (2014). An integrated approach reveals regulatory controls on bacterial translation elongation. Cell *159*, 1200–1211, 10.1016/j.cell.2014.10.043.

Tadros, W., and Lipshitz, H.D. (2005). Setting the stage for development: mRNA translation and stability during oocyte maturation and egg activation in Drosophila. Dev. Dyn. *232*, 593–608, 10.1002/dvdy.20297.

Tanaka, T., Kato, Y., Matsuda, K., Hanyu-Nakamura, K., and Nakamura, A. (2011). Drosophila Mon2 couples Oskar-induced endocytosis with actin remodeling for cortical anchorage of the germ plasm. Development *138*, 2523–2532, 10.1242/dev.062208.

Travers, K.J., Patil, C.K., Wodicka, L., Lockhart, D.J., Weissman, J.S., and Walter, P. (2000). Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell *101*, 249–258, 10.1016/S0092-8674(00)80835-1.

Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell *141*, 344–354, 10.1016/j.cell.2010.03.031.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419, 10.1126/science.1260419.

Vanzo, N.F., and Ephrussi, A. (2002). Oskar anchoring restricts pole plasm formation to the posterior of the Drosophila oocyte. Development *129*, 3705–3714.

Varanasi, U.S., Klis, M., Mikesell, P.B., and Trumbly, R.J. (1996). The Cyc8 (Ssn6)-Tup1 corepressor complex is composed of one Cyc8 and four Tup1 subunits. Mol. Cell. Biol. *16*, 6707–6714, 10.1128/MCB.16.12.6707.

Walter, P., and Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. Science *334*, 1081–1086, 10.1126/science.1209038.

Walters, R.W., Muhlrad, D., Garcia, J., and Parker, R. (2015). Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in Saccharomyces cerevisiae. RNA *21*, 1660–1671, 10.1261/rna.053116.115.

Wang, Q., Zhang, Z., Blackwell, K., and Carmichael, G.G. (2005). Vigilins bind to promiscuously A-to-I-edited RNAs and are involved in the formation of heterochromatin. Curr. Biol. *15*, 384–391, 10.1016/j.cub.2005.01.046.

Webb, S., Hector, R.D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol. *15*, R8, 10.1186/gb-2014-15-1-r8.

Weber, V., Wernitznig, A., Hager, G., Harata, M., Frank, P., and Wintersberger, U. (1997). Purification and nucleic-acid-binding properties of a Saccharomyces cerevisiae protein involved in the control of ploidy. Eur. J. Biochem. *249*, 309–317.

Webster, M.W., Chen, Y.-H., Stowell, J.A., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J., and Passmore, L.A. (2018). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Mol. Cell *70*, 1089–1100.e8, 10.1016/j.molcel.2018.05.033.

Weidner, J., Wang, C., Prescianotto-Baschong, C., Estrada, A.F., and Spang, A. (2014). The polysome-associated proteins Scp160 and Bfr1 prevent P body formation under normal growth conditions. J. Cell Sci. *127*, 1992–2004, 10.1242/jcs.142083.

Wen, W.-L., Stevenson, A.L., Wang, C.-Y., Chen, H.-J., Kearsey, S.E., Norbury, C.J., Watt, S., Bähler, J., and Wang, S.-W. (2010). Vgl1, a multi-KH domain protein, is a novel component of the fission yeast stress granules required for cell survival under thermal stress. Nucleic Acids Res. *38*, 6555–6566, 10.1093/nar/gkq555.

Wilson, M.A. (2014). Metabolic role for yeast DJ-1 superfamily proteins. Proc. Natl. Acad. Sci. U. S. A. *111*, 6858–6859, 10.1073/pnas.1405511111.

Wintersberger, U., Kühne, C., and Karwan, A. (1995). Scp160p a new yeast protein associated with the nuclear membrane and the endoplasmic reticulum, is necessary for maintenance of exact ploidy. Yeast *11*, 929–944, 10.1002/yea.320111004.

Woerner, A.C., Frottin, F., Hornburg, D., Feng, L.R., Meissner, F., Patra, M., Tatzelt, J., Mann, M., Winklhofer, K.F., Hartl, F.U., et al. (2016). Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. Science *351*, 173–176, 10.1126/science.aad2033.

Yang, J., Hao, X., Cao, X., Liu, B., and Nyström, T. (2016). Spatial sequestration and detoxification of Huntingtin by the ribosome quality control complex. Elife *5*, 1–14, 10.7554/eLife.11792.

Yonashiro, R., Tahara, E.B., Bengtson, M.H., Khokhrina, M., Lorenz, H., Chen, K.-C., Kigoshi-Tansho, Y., Savas, J.N., Yates, J.R., Kay, S.A., et al. (2016). The Rqc2/Tae2

subunit of the ribosome-associated quality control (RQC) complex marks ribosome-stalled nascent polypeptide chains for aggregation. Elife *5*, e11794, 10.7554/eLife.11794.

Yoon, J.-H., Abdelmohsen, K., and Gorospe, M. (2013). Posttranscriptional gene regulation by long noncoding RNA. J. Mol. Biol. *425*, 3723–3730, 10.1016/j.jmb.2012.11.024.

Yu, C.-H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M.S., and Liu, Y. (2015). Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. Mol. Cell *59*, 744–754, 10.1016/j.molcel.2015.07.018.

Yu, X., Liu, Q., He, J., Huang, Y., Jiang, L., Xie, X., Liu, J., Chen, L., Wei, L., and Qin, Y. (2018). Vigilin interacts with CTCF and is involved in the maintenance of imprinting of IGF2 through a novel RNA-mediated mechanism. Int. J. Biol. Macromol. *108*, 515–522, 10.1016/j.ijbiomac.2017.11.109.

Yue, Y., Liu, J., and He, C. (2015). RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. Genes Dev. *29*, 1343–1355, 10.1101/gad.262766.115.

Zabinsky, R.A., Weum, B.M., Cui, M., and Han, M. (2017). RNA Binding Protein Vigilin Collaborates with miRNAs To Regulate Gene Expression for Caenorhabditis elegans Larval Development. G3 (Bethesda). *7*, 2511–2518, 10.1534/g3.117.043414.

Zhang, H., Elbaum-Garfinkle, S., Langdon, E.M., Taylor, N., Occhipinti, P., Bridges, A.A., Brangwynne, C.P., and Gladfelter, A.S. (2015). RNA Controls PolyQ Protein Phase Transitions. Mol. Cell *60*, 220–230, 10.1016/j.molcel.2015.09.017.

Zhou, J., Wang, Q., Chen, L.-L., and Carmichael, G.G. (2008). On the mechanism of induction of heterochromatin by the RNA-binding protein vigilin. RNA *14*, 1773–1781, 10.1261/rna.1036308.

Zhou, M., Guo, J., Cha, J., Chae, M., Chen, S., Barral, J.M., Sachs, M.S., and Liu, Y. (2013). Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature *495*, 111–115, 10.1038/nature11833.

7. LIST OF PUBLICATIONS

Review Article:

Cheng, M.H., and Jansen, R.-P. (2017). A jack of all trades: the RNA-binding protein vigilin. Wiley Interdiscip. Rev. RNA *8*, e1448, 10.1002/wrna.1448.

Research Article:

Cheng, M.H., Hoffmann, P.C., Franz-Wachtel, M., Sparn, C., Seng, C., Maček, B., and Jansen, R.-P. (2018). The RNA-Binding Protein Scp160p Facilitates Aggregation of Many Endogenous Q/N-Rich Proteins. Cell Rep. *24*, 20–26, 10.1016/j.celrep.2018.06.015.

8. PERSONAL CONTRIBUTIONS

1. For the Review Article **"A jack of all trades: The RNA-binding protein vigilin"**, I performed:
   - Literature research
   - Preparation, writing, and editing of the manuscript
2. For the Research Article **"The RNA-binding protein Scp160p facilitates the aggregation of many endogenous Q/N-rich proteins"**, I contributed to the:
   - Conceptualization of the project
   - Experimental design, performance, and analyses
   - Optimization, performance, and analysis of the combined filter-trap binding and quantitative mass spectrometry
   - Preparation, writing, and editing of the manuscript

9. APPENDIX

9.1 – Review Article – A jack of all trades: the RNA-binding protein vigilin.

(Cheng and Jansen, 2017)
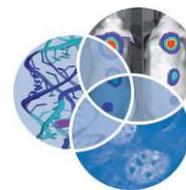
A jack of all trades: the RNA-binding protein vigilin

Matthew HK Cheng ⓘ[1,2] and Ralf-Peter Jansen ⓘ[2]*

The vigilin family of proteins is evolutionarily conserved from yeast to humans and characterized by the proteins' 14 or 15 hnRNP K homology (KH) domains, typically associated with RNA-binding. Vigilin is the largest RNA-binding protein (RBP) in the KH domain-containing family and one of the largest RBP known to date. Since its identification 30 years ago, vigilin has been shown to bind over 700 mRNAs and has been associated with cancer progression and cardiovascular disease. We provide a brief historic overview of vigilin research and outline the proteins' different functions, focusing on maintenance of genome ploidy, heterochromatin formation, RNA export, as well as regulation of translation, mRNA transport, and mRNA stability. The multitude of associated functions is reflected by the large number of identified interaction partners, ranging from tRNAs, mRNAs, ribosomes and ribosome-associated proteins, to histone methyltransferases and DNA-dependent protein kinases. Most of these partners bind to vigilin's carboxyterminus, and the two most C-terminal KH domains of the protein, KH13 and KH14, represent the main mRNA-binding interface. Since the nuclear functions of vigilins in particular are not conserved, we outline a model for the basal functions of vigilins, as well as those which were acquired during the transition from unicellular organisms to metazoa. © 2017 Wiley Periodicals, Inc.

GETTING TO KNOW JACK: AN INTRODUCTION

Control of gene expression is, to an extensive level, mediated by RNA-binding proteins (RBPs). These proteins recognize and bind *cis*-sequences and/or secondary structures within target RNAs, thereby regulating their localization, translation, and/or stability. A single RBP may be able to bind and regulate multiple mRNA species, and thus their influence over cell function is extensive.[1] A striking example of such proteins is the vigilin proteins. These proteins constitute a family of evolutionarily conserved RBPs characterized by their tandem arrangement of multiple heterogeneous

ribonucleoprotein particle (hnRNP) K homology (KH) domains (Box 1). Their vast influence over cellular processes is demonstrated by their large number of potential mRNA targets (over 700)[1,2] and their link to diseases such as hepatocellular carcinoma,[3] breast cancer,[4] and obesity in mice.[2] Accordingly, knockdown or deletion of vigilin in human cell lines or *Saccharomyces cerevisiae* leads to decreased viability and fitness, respectively.[5,6]

The most striking consequences upon knockdown or deletion of vigilin are defects in chromosome and ploidy maintenance. Deletion of vigilin in *S. cerevisiae* resulted in mutants with increased DNA content/ploidy[5,17] whereas knockdown of vigilin in *Drosophila* S2 and human liver cells led to defects in chromosome condensation manifesting in chromosome misalignment and lagging during division.[18–20] While these observations attribute a role for vigilin in chromatin organization, vigilin has also been observed to influence many stages of RNA metabolism, including mRNA transport,[21,22]

---
*Correspondence to: ralf.jansen@uni-tuebingen.de

[1]International Max Planck Research School, Tuebingen, Germany

[2]Interfaculty Institute of Biochemistry, Tuebingen, Germany

Conflict of interest: The authors have declared no conflicts of interest for this article.

THE hnRNP K HOMOLOGY (KH) DOMAIN

KH domains were characterized based on their homology to the human hnRNP K.[7] Although there are two types of KH domains—typical in eukaryotic and prokaryotic proteins, respectively—they share a common 'minimal' $\beta\alpha\alpha\beta$ structure (a).[8] The GXXG motif that defines a classical KH domain is found in the loop between the alpha helices. In the eukaryotic type I KH domain, there is a C-terminal extension consisting of a variable loop, β strand, and α helix (a). The overall fold of a type I KH domain involves an antiparallel sheet formed by the β strands sitting adjacent to the three α helices (b). A hydrophobic groove is formed on one side by the α-GXXG-α helix-loop-helix, and on the other side by the $\beta2$ strand and variable loop (b). This groove accommodates a nucleic acid tetramer whose sequence is specified by hydrogen bonding to amino acid side chains.[8,9] Structural data from KH domains of other proteins show that the hydrogen bond interactions facilitate adenines and cytosines, but typically not guanines.[9] While the structure of KH14 of the human vigilin[10] is shown in (b), the structures of KH domains 1, 4, 6, 8, 12, and 13 of human vigilin have also been deposited as PDB files.[11–16]

PDB ID: **2CTM** (KH14) and **1J5K** (NA)
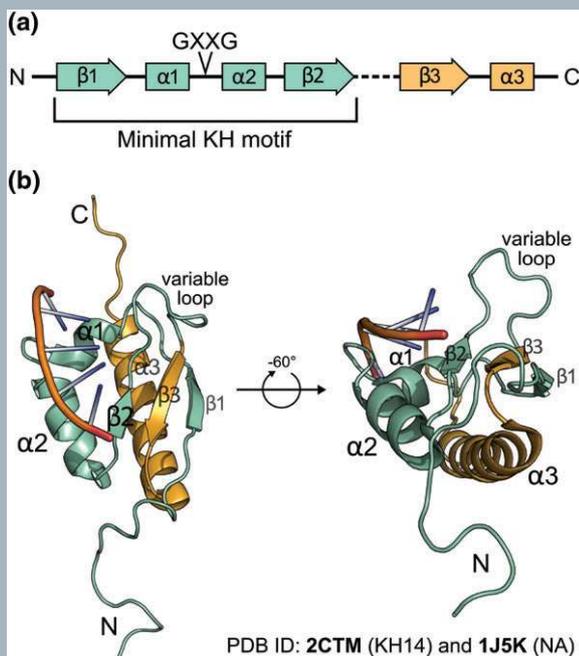
translation,[2,23–25] degradation,[4,26] and formation of stress granules[27] and processing bodies (P-bodies).[28]

Despite an abundance of research into the vigilin proteins since their identification 30 years ago, our understanding of their exact functions in the cell remains unclear. This review will bring together the major findings on the different vigilin homologs, with a focus on the molecular basis of their functions. By doing so, we aim to highlight parallels and differences between homologs in their many roles. Such a consideration of vigilins' cellular functions will help to consolidate and strengthen our understanding of this family of protein.

WHAT IS IN A NAME? A BRIEF HISTORY OF THE VIGILIN PROTEIN FAMILY

Through 30 years of vigilin research, the members of this protein family have been given various names, all of which remain in use today. While the continued use of various names create ambiguity among the research on the homologs, they also reflect the many functions which have been attributed to vigilin. The human homolog was the first to be described in 1987[29] (Figure 1(a)). Unexpectedly, it was identified due to its high-density lipoprotein (HDL) binding activity, and not nucleic acid-binding. This led to the name HDLBP (high-density lipoprotein binding protein). The cDNA and amino acid sequences of chicken and human vigilin/HDLBP were published 5 years later in 1992.[30,31] Since the authors were unaware that it had been previously named HDLBP, they named the protein 'vigilin' after a conserved VIG (valine-isoleucine-glycine) motif in the N-terminal region of the chicken homolog. In both proteins, 14 tandem repeats were identified, which would not be realized as the nucleic acid-binding hnRNP KH domains (Box 1) until 1993.[7]

The publication of the human and chicken vigilin primary structures eventually led to the hesitant realization that the two proteins—viewed to be unrelated at the time—are homologs. The hesitation was largely driven by the discrepancy between the proposed role of HDLBP as a cell surface HDL receptor and the subsequent cytoplasmic staining observed by researchers studying the 'human homolog' of chicken vigilin.[29–32] However, the high degree of homology and the cross-reactivity of both an anti-vigilin antibody and Northern RNA probes reconciled the two as homologs.[31,33]

In 1995, the next vigilin gene was 'accidentally' isolated from the budding yeast *S. cerevisiae* while screening a yeast DNA expression library with antiserum raised against a protein fraction from a

(a)

(b)

Organism	Gene name	Localization	Function(s)	References
H. sapiens	HDLBP	Nucleus, cytoplasm, and cytoplasmic side of ER	HDL-binding	19
			Nuclear-cytoplasmic tRNA shuttling	45,46
			Ribosome association	34
			Heterochromatin regulation	29
			Gene imprinting	55
			Translation regulation	67
M. musculus	Hdlbp		Translation regulation	2
G. gallus	Vigilin	?	?	
X. laevis	hdlbp	?	mRNA stabilization	16
S. cerevisiae	SCP160	Cytoplasm and ER	Ploidy maintenance	5,7
			Polysome association	38,39
			Pheromone sensing and mating	12,80
			Translation regulation	14
			mRNA localization	12
			Translation elongation	15
			P-body regulation	18
D. melanogaster	DDP1	Cytoplasm and ER	Heterochromatin regulation	8,28
			Translation regulation	13
S. pombe	vgl1	Cytoplasm and ER	Stress granule component	17

FIGURE 1 | Timeline of the identification of vigilin homologs and their functions. (a) Timeline highlighting the years in which the vigilin homologs were identified. (b) A table summarizing the reported localization and functions of the vigilin homologs.

ribonuclease H purification.[5] Immunofluorescence microscopy revealed that the protein is localized to the nuclear envelope and endoplasmic reticulum (ER). Deletion mutants lacking the gene had reduced viability and showed an increase in ploidy, giving rise to yet another name for vigilin: Scp160p (an apparent 160 kDa <u>S</u>. *cerevisiae* protein <u>c</u>ontrolling the <u>p</u>loidy).[5] Shortly after the characterization of KH domains and their nucleic acid-binding capability, both the human and *S. cerevisiae* vigilin proteins were reported to bind nucleic acid.[17,34] While the human vigilin was found to bind tRNAs,[34] the *S. cerevisiae* vigilin was reported to bind both DNA and RNA.[17]

At around the same time, the *Xenopus laevis* homolog was identified as an estrogen-inducible protein which bound specifically to the 3′ untranslated region (UTR) of vitellogenin mRNA.[35,36] In contrast to that, the *Drosophila* vigilin was characterized as a protein binding single-stranded dodeca-satellite DNA, a highly repetitive element found in the pericentric regions of chromosomes.[37] Thus, the *Drosophila* homolog was given a fourth name, DDP1 (*Drosophila* <u>d</u>odeca-satellite-binding protein 1). Its binding at the pericentric

region shifted research focus to vigilin as a chromosome interactor necessary to maintain ploidy through division.

The latest vigilin homolog identified is from the fission yeast *Schizosaccharomyces pombe*.[27] Interestingly, deletion of *S. pombe* vigilin—named Vgl1 in the study—did not influence ploidy, in contrast to the phenotype reported for *S. cerevisiae*, *Drosophila*, and mammals.[5,19,27,38] Instead, *S. pombe* vigilin was implicated in stress response as a novel component of stress granules.[27]

THE SUM OF ITS PARTS: DOMAIN ARCHITECTURE OF THE VIGILIN PROTEINS

The vigilin family of proteins is characterized by its domain architecture, which consists of multiple tandemly arranged KH domains (Figure 2(a)). KH domains are one of the best characterized family of nucleic acid-binding domains (Box 1)[7] and one typical feature of KH domains is the presence of a
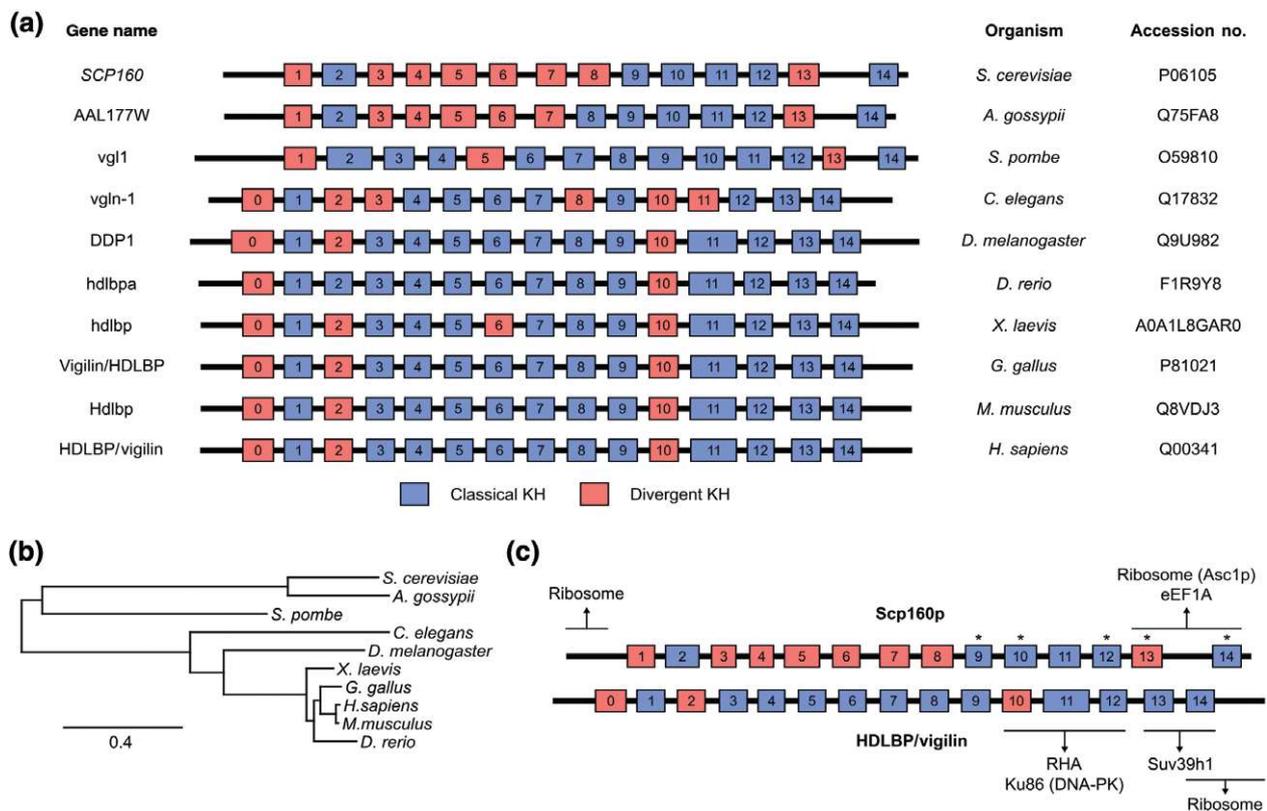
FIGURE 2 | The domain arrangement of vigilin homologs. (a) Schematics of vigilin homologs are shown with classical KH domains in blue and divergent KH domains in red. All schematics are aligned based on KH1. The vigilins of lower eukaryotes (with only 14 KH domains) are shown at the top. (b) The evolutionary relationship between the various members of the vigilin protein family. The phylogenetic tree is generated based on multiple sequence alignment of the vigilin proteins.[39] Branch length represent amount of genetic change. (c) Schematic of the *S. cerevisiae* and *H. sapiens* vigilins highlighting the known vigilin domains which interact with other proteins. Asterisks (*) indicates the KH domains of *S. cerevisiae* Scp160p with RNA-contacting amino acids. For more information, see main text.

GXXG motif that contributes to nucleic acid-binding.[8,9] This motif may be altered or lost in some KH domains, giving rise to so-called divergent/degenerate KH domains, in contrast to classical KH domains.[8]

Despite vigilin being initially reported to contain 14 KH domains, subsequent publications have reported an N-terminal 15th KH domain (numbered '0').[6,40,41] The confusion in numbers likely stems from the fact that KH0 is divergent and could have been missed when the initial reports discovered vigilin's KH repeats.[30,31] Nevertheless, secondary structure prediction and multiple sequence alignment of homologs from higher eukaryotes reveal the presence of 15 KH domains (Figure 2(a)). In contrast, the *S. cerevisiae* and *S. pombe* homologs only contain 14 KH domains.[17,27]

Vigilin's complement of RNA-binding domains is exceptional and makes it the largest protein in the KH family.[2,42] Despite the conservation of their general architecture, the vigilins differ in their number and order of classical and divergent KH domains[8,17,43] (Figure 2(a)). For example, in the human homolog,

12 out of 15 KH domains are classical whereas only 8 out of 14 in the *S. cerevisiae* homolog are classical.[17,42] Hitherto, few studies have investigated whether the variation in ratios of classical to divergent KH domains and/or their arrangement can lead to functional differences and reflect different binding potential for RNA. While *Drosophila* vigilin (9 out of 15 KH domains being classical) can complement *S. cerevisiae* mutants lacking the yeast homolog and maintain exact ploidy,[37] *S. pombe* vigilin (with a domain composition similar to *Drosophila*) is dispensable for ploidy maintenance in the fission yeast.[27] Thus, the assortment of classical and divergent KH domains may not correlate to function, although this remains to be addressed directly, for example, by structure-function analyses or domain swapping.

A number of studies have shown the importance of the C-terminus of the vigilin proteins. Most recently, RNA-contacting sites were mapped to be spread along the C-terminal half of *S. cerevisiae* vigilin, where most classical KH domains are located

(Figure 2(c) and Box 2).[44] This finding is in agreement with previous experiments demonstrating that truncation of the C-terminal most KH13-14 domains abolishes RNA-binding[22,25] and polysome association (see below).[41,45,46] The sequence C-terminal of KH14 in human vigilin has also been implicated in mediating ribosome interaction (see below).[41]

Interestingly, the C-terminal KH13-14 domains of human vigilin also mediates interaction with the histone methyltransferase Suv39h1 (see below).[19] Furthermore, KH10-12 mediates interactions with a number of proteins involved in the nonhomologous end joining (NHEJ) pathway of DNA double-stranded break repair and RNA editing.[19,48] Together, these studies highlight a special importance of vigilin's C-terminal half for its cytoplasmic and nuclear functions.

In addition to the KH domains, the vigilin proteins also contain a putative nuclear localization signal (NLS) between KH3 and KH4, and a putative nuclear export signal (NES) in its N-terminus.[17,43,49]

However, these signals might not be functional in all family members, underscoring the functional discrepancy between homologs. In a nuclear import reporter assay, the NLS from human vigilin is functional.[49] In contrast, mutation of the putative NLS in S. cerevisiae vigilin did not produce any observable growth defects, and inhibiting nuclear export pathways failed to trap the protein in the nucleus.[43] The observed difference in the proteins' localizations, however, may be due to the different experimental approaches and it cannot be excluded that vigilin homologs of higher eukaryotes or perhaps only the human homolog, have evolved to play a more prominent role in the nucleus (Figures 2(b) and 3(b)).

SUBCELLULAR DISTRIBUTION OF VIGILIN PROTEINS

As vigilin was initially believed to be a receptor for HDL, it was hypothesized to be present at the cell

---

BOX 2

RNA-BINDING BY VIGILINS

Vigilins, with their multiple KH domains, bind nucleic acids promiscuously. In addition to the nucleic acid species mentioned in text, photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) analysis showed that mouse vigilin can also bind a variety of RNA species including transfer RNAs, ribosomal RNAs, small nuclear and small nucleolar RNAs, long intergenic noncoding RNAs, microRNAs, and mitochondrial RNAs.[2] In vitro genetic selection using Xenopus vigilin identified the consensus binding motif as $(A)_nCU$ and $CU(A)_n$ within a 75 nt long single-stranded region.[47] In silico analysis of microarray data from S. cerevisiae showed UGAAAAAUUUU as the consensus motif, although with low confidence.[1] Recent analysis of PAR-CLIP data from mice revealed the vigilin-binding motif to be a tandem of CHHC or CHYC (H = A/C/U and Y = C/U) separated by two, five, or eight nucleotides.[2] Although the latter motifs might indeed represent the real vigilin-binding site, it is obvious from the three studies that there is a lack of sequence specificity. Notably, this lack of sequence specificity in addition to the preference of vigilin for single-stranded nucleic acid suggests that recognition may be largely mediated by the absence of secondary structure.

UV cross-linking and mass spectrometry has mapped the RNA-contacting sites in the S. cerevisiae vigilin.[44] Seven contact sites were mapped that are spread along vigilin's C-terminal half where its classical KH domains are concentrated. In agreement with the importance of KH13-14 for RNA-binding,[22,25] three of seven of the mapped sites lie in the last two KH domains of the S. cerevisiae vigilin.

Consensus Vigilin Binding Sequences			
Homolog	Consensus Sequence	Experimental Method	Reference
X. laevis	$(A)_nCU$ and $CU(A)_n$	In vitro genetic selection	47
S. cerevisiae	UGAAAAAUUUU	RIP and microarray	1
M. musculus	CHHCNNCHYC	PAR-CLIP	2
	CHHCNNNNNCHYC		
	CHHCNNNNNNNNCHYC		
	(H = A/C/U and Y = C/U)		

---

surface.[29,30] Subsequent biochemical and microscopic data showed that the majority of vigilin is actually cytoplasmic.[5,32,50] Moreover, a fraction of vigilin is enriched on the cytoplasmic side of the outer nuclear membrane and rough ER, but not other organelles.[5,27,50–52] Since vigilins bind to numerous mRNAs encoding membrane or secreted proteins,[2,25] which are generally translated at the ER, this finding is in good agreement with vigilin's reported polysome association and role in translation.[50,52] In *S. pombe*, in response to thermal, osmotic, nutritional, and arsenite stresses, vigilin re-localizes from the ER surface into cytoplasmic RNA granules suggested to be stress granules.[27] Re-localization to RNA granules upon thermal stress has also been observed in *S. cerevisiae*, but whether these are stress granules or P-bodies was not investigated.[27] Indeed, despite vigilin's importance in regulating P-body formation in *S. cerevisiae*,[28] it is unclear if vigilin itself is a P-body component.

Although vigilin is absent from the nucleus in budding and fission yeasts,[27,43,51,52] a small fraction is present in the nucleus in higher eukaryotes.[19,49,50,53] Here, it is present on heterochromatin but not euchromatin,[50] in support of its proposed role in heterochromatin formation.

FUNCTION(S) OF VIGILIN IN THE NUCLEUS

Vigilin as a Nucleocytoplasmic Shuttle of tRNAs

Vigilin's presence in the nucleus and cytoplasm, combined with the existence of a functional NLS in the human protein suggest that mammalian vigilins might shuttle between the nucleus and the cytoplasm. In agreement, vigilin has been found in two distinct but similar multiprotein complexes in the nucleus and cytoplasm of human epithelial cells.[53,54] Besides vigilin, both complexes (named VCC$_N$ and VCC$_C$ for nuclear and cytoplasmic vigilin-containing complexes) contain tRNA and the translation elongation factor EF-1α,[54] which raised the possibility for vigilin to be
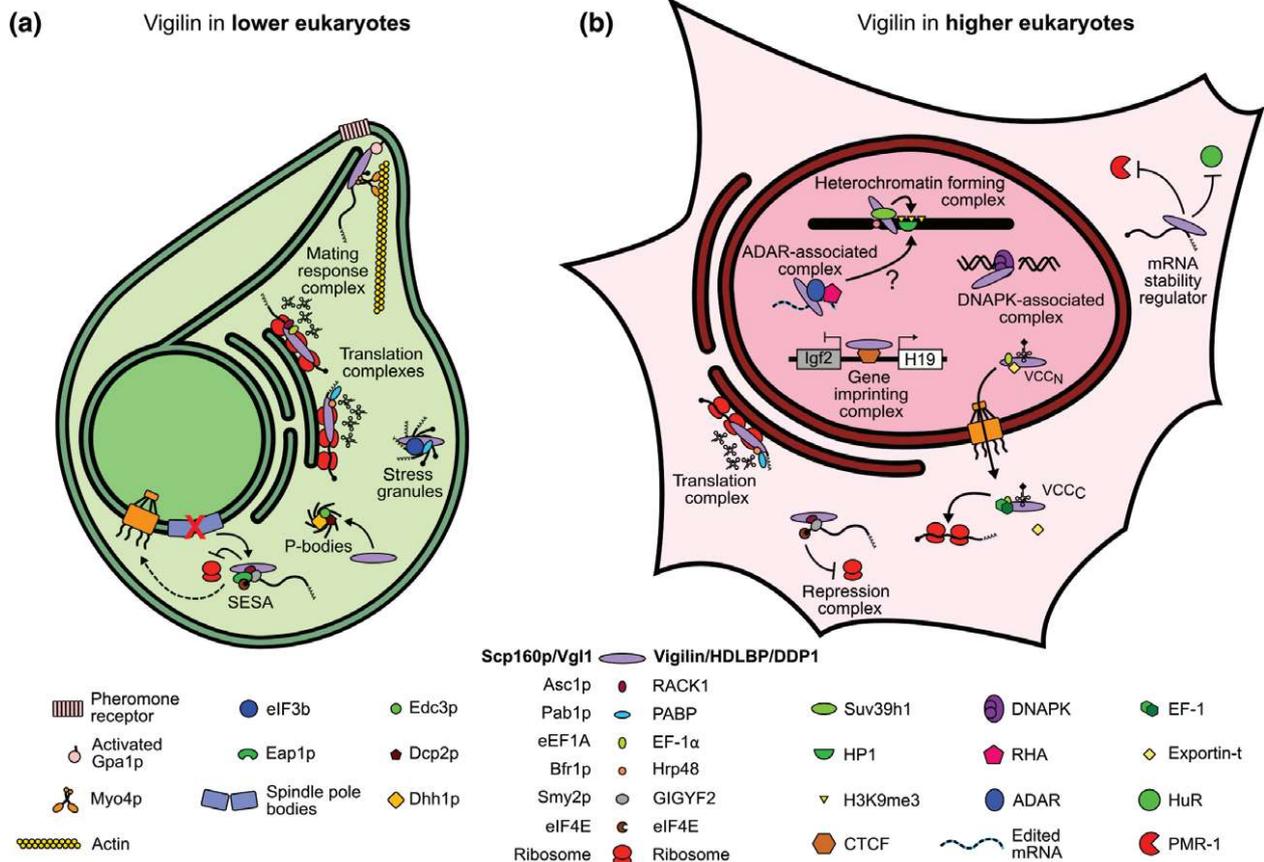

FIGURE 3 | Models of vigilin functions. Vigilin-associated complexes are shown in (a) lower eukaryotes (represented by a shmooing *S. cerevisiae* cell) and in (b) higher eukaryotes (represented by a HeLa cell). Proteins with homologous roles are shown in the middle list. Proteins with roles specific to either lower or higher eukaryotes are listed under the respective cell.

involved in tRNA nuclear export. Indeed, the tRNA-specific nuclear export receptor exportin-t is a component of VCC$_N$ but not VCC$_C$, and interacts physically with vigilin.[53] Furthermore, both purified VCC$_N$ and recombinant vigilin was able to mediate nuclear export of tRNAs in co-injection assays.[53]

In the cytoplasm, VCC$_N$ converts to VCC$_C$ by release of exportin-t and association with translation elongation factor EF-1,[53] which led to the model whereby vigilin—as part of a complex—exports tRNAs into the cytoplasm and subsequently delivers them to the translation machinery (Figure 3(b)).[53] This model is in agreement with later work in *S. cerevisiae* which attribute a role for the yeast homolog in tRNA recycling (see below).[25]

In both VCC$_N$ and VCC$_C$, tRNA-binding is mediated by vigilin.[54] Moreover, binding is specific to tRNAs as purified nucleic acid-free VCC did not bind other RNA species when provided human total RNA.[34] Interestingly, excess tRNA could not compete with vitellogenin mRNA for binding to *Xenopus* vigilin (see below). This suggests that vigilins from distinct species might have different affinities for different RNA species.

Vigilin in Chromosome Biology

Loss of vigilin gives rise to chromosomal defects and increase in ploidy. Furthermore, in *S. cerevisiae* vigilin mutants, gene silencing at telomeres and the mating-type locus is lost.[5,17,55] However, the precise mechanism(s) through which vigilin acts on chromosome and ploidy maintenance remains unclear. The observed exclusion of *S. cerevisiae* and *S. pombe* vigilins[27,43,55] from the nucleus puts into question whether they act directly on chromatin. Indeed, *S. cerevisiae* vigilin has been proposed to maintain ploidy by repressing translation of *POM34* mRNA and might act as a quality control checkpoint in spindle pole body (SPB) duplication (Figure 3(a) and see below).[24] However, it is conceivable that vigilins in metazoans have evolved to act more directly on chromatin. Several studies suggest that metazoan vigilin could facilitate formation of heterochromatin by recruiting chromatin remodeling factors to heterochromatic sites. In both mammals and *Drosophila*, vigilin has been shown to bind the highly repetitive DNA sequences flanking centromeres (satellite DNA) typically kept in a heterochromatic state.[19,37,38] More specifically, the *Drosophila* vigilin binds the C-rich strand of the dodeca-satellite but not the complementary G-rich strand or the duplex *in vitro*.[37] In addition, human vigilin co-immunoprecipitates with alpha- and beta-satellites.[38] It is interesting to note

that this DNA-binding preference is reflected in vigilin's RNA-binding consensus motifs, which are unstructured, single-stranded, and G-poor (Box 2).[1,2,47] Co-immunoprecipitation and co-localization data suggest that vigilin recruits the histone methyltransferase Suv39h1 and heterochromatin protein 1 (HP1) to heterochromatic regions of DNA.[19,38,56] The interaction between the *Drosophila* vigilin and HP1 is RNA dependent and the two proteins form a nuclear complex with two other hnRNPs.[56] The recruitment of Suv39h1 leads to the trimethylation of lysine 9 (K9) in the histone 3 (H3) tail, a modification that is recognized by HP1, which recruits additional histone modification factors.[57–60] Indeed, polytene chromosomes from *vigilin$^{-/-}$* flies show reduced H3K9me3 compared to those from wild-type flies,[52] and knockdown of vigilin in *Drosophila* S2 cells resulted in mislocalization of HP1.[38]

Strikingly, ectopic expression of the vigilin KH13-14 domain that interacts with Suv39h1 in HEK293 cells led to a defect in chromosome segregation and disrupted beta-satellite heterochromatin.[19] The authors of this study argued—but did not test directly—that this is due to competition with the endogenous vigilin for and sequestration of the histone methyltransferase. However, the expression of a KH10-14 fragment did not mimic this effect.[19] This might be due to the observed higher abundance of KH13-14 in the nucleus, which may better compete for and sequester Suv39h1 away from endogenous vigilin. In support of this, the KH13-14 fragment co-immunoprecipitates with Suv39h1 better than the KH10-14 fragment.[19]

In addition to the interaction with histone methyltransferase, vigilin can also bind to DNA-dependent protein kinase (DNA-PK) via its KH10-12 domains, implicating a role in maintenance of DNA fidelity.[38] Aside from DNA-PK, KH10-12 also mediates interaction with another nuclear protein complex containing the DExD/H-box helicase RNA helicase A (RHA/DHX9) and the RNA-editing enzyme ADAR (<u>a</u>denosine <u>d</u>eaminase <u>a</u>cting on ds<u>RN</u>A).[19,38] Vigilin can bind RNAs that are edited by ADAR,[38] prompting the authors to suggest the vigilin-RHA-ADAR complex as an alternative pathway to RNA interference (RNAi) in heterochromatin formation, although this remains untested.

These data strongly support nuclear functions of vigilin in metazoa, again suggesting the evolution of additional functions for vigilin. Although there is little evidence for a nuclear form of vigilin in *S. cerevisiae*, its loss is accompanied by reduced telomere association of the histone-binding protein Sir3p, a component of the chromatin silencing complex that is involved in

silencing of telomeric subregions.[55,61] However, whether this reflects a direct or indirect effect is unknown. Complementation experiments using the various vigilin homologs might shed light into how the mechanism of vigilin function in chromatin maintenance may have diverged through evolution.

A yeast two hybrid screen revealed an interaction between human vigilin and the CCCTC-binding factor (CTCF).[62] CTCF, among other functions, is involved in gene imprinting, a process whereby gene expression from parental alleles are reciprocally regulated in a DNA methylation-dependent manner.[62,63] In human and mice, two examples of imprinted genes are insulin-like growth factor (*Igf2*) and *H19*. Strikingly, vigilin binds the *H19* differentially methylated region (DMR) in a CTCF-dependent manner.[62] Up- or downregulation of vigilin led to changes in *Igf2* and *H19* expression which parallel those observed when CTCF is up- or downregulated. This led the authors to propose that vigilin and CTCF function together in gene expression regulation.[62] However, it remains to be shown whether vigilin's association with CTCF and the DMR is direct or mediated by nuclear RNAs (e.g., long noncoding RNAs), and how vigilin participates in CTCF-mediated gene imprinting.

VIGILIN'S ROLE IN THE CYTOPLASM

Translation Stimulation by Vigilin

Vigilin expression in a variety of cell lines correlate with demand for protein synthesis. For example, transcription of vigilin is upregulated in proliferating but not resting chondrocytes.[33] In rats, vigilin protein levels are increased in pancreatic extracts of fed animals, where trypsin production is stimulated, compared to starved animals.[40] Moreover, vigilin is overexpressed in hepatocellular carcinomas and knockdown of vigilin in these cells suppresses proliferation and mobility.[3] These observations together with vigilin's association with free ribosomes and polysomes at the rough ER suggested a potential role in translation.[32,45,50–52,64] Recently, Mobin et al. found that vigilin promotes translation of a number of secreted liver proteins, including apolipoprotein B, thereby modulating the blood plasma lipid profile of mice.[2] Interestingly, the mRNAs whose translation state changes upon loss of vigilin in *S. cerevisiae* are enriched for those encoding extracellular, cell wall, and ER-to-Golgi transport proteins.[25] Thus, vigilins may bind and influence the translation of mRNAs encoding membrane-associated and secreted proteins.

Yeast vigilin might associate with polysomes by two different but not exclusive modes. In one scenario, vigilin is part of a >1.3 MDa complex that also contains poly(A) binding protein (Pab1p) and the RBP Bfr1p,[65] and the association of this complex with polysomes is dependent on the latter.[64] Interestingly, a similar complex in *Drosophila* embryos—also containing vigilin and PABP—was identified to enhance translation of mRNAs upon binding to a *cis*-element in their 3′UTR.[23] In this case, the hnRNP A/B family protein Hrp48 was found as the RBP in the complex. However, despite their similar makeup, the two vigilin-containing complexes from *S. cerevisiae* and *Drosophila* might not be functionally equivalent. Whereas the *Drosophila* complex is resistant to RNase, the yeast complex is not.[23,65] Furthermore, it is unknown whether the *Drosophila* complex associates with polysomes like the yeast complex.

In the other scenario, vigilin associates with ribosomes in close proximity to Asc1p/RACK1 (receptor of activated protein C kinase in mammals), which is an integral component of the small ribosomal subunit. Vigilin also interacts at ribosomes with the elongation factor eEF1A, as identified by co-immunoprecipitation after chemical cross-linking.[45] Yeast vigilin's interaction with Asc1p/RACK is dependent on KH13-14 (Figure 2(c)), and loss of either Asc1p/RACK1 or KH13-14 diminishes vigilin's polysome association.[45] Since vigilin's association with polysomes is RNA-dependent[65] and truncation of KH13-14 also abolishes vigilin binding to several target mRNAs,[25] KH13-14 could mediate polysome association via mRNA-binding. A point mutation analysis is needed to distinguish whether vigilin's polysome association is mediated by RNA-binding, by interaction with Asc1p/RACK1, or both. The availability of KH domain mutants impaired for nucleic acid-binding[66] as well as the identification of the amino acids with RNA contacts[44] will allow an informed and precise approach to this analysis.

The inclusion of eEF1A in the vigilin-Asc1p/RACK1 complex is reminiscent of the tRNA shuttling complex described earlier[53] and may link tRNA-binding/transport with translation. In support of this, loss of vigilin in *S. cerevisiae* leads to ribosomal pausing on a subset of mRNAs whose codons are arranged to minimize switching of cognate tRNA species between successive occurrences of the same amino acid (a phenomenon termed autocorrelation).[25,67] Autocorrelation thereby enhances translation efficiency by allowing substrate channeling of tRNAs.[67,68] Based on the high autocorrelation of mRNAs whose translation is slowed in yeast mutants lacking vigilin, it is reasonable to expect that cognate

tRNAs for autocorrelated codons would be most depleted from ribosomes in vigilin mutants. Surprisingly in the absence of vigilin, the tRNAs most depleted from ribosomes are those decoding anticorrelated codons (where successive occurrences of the same amino acid uses alternating synonymous codons) within vigilin's mRNA targets. Therefore, vigilin might function to complement autocorrelation and substrate channeling by limiting loss of deacylated tRNAs decoding anticorrelated codons from the vicinity of translating polysomes (Figure 3(a)).[25] This model is also consistent with PAR-CLIP data from mouse liver showing enrichment of vigilin-binding sites within the coding region of target mRNAs.[2] Finally, this model reconciles the high throughput interaction data showing contact between yeast vigilin and ribosomal proteins present at both the proximal and distal faces of the ribosome.[69] This coincides with a role for vigilin in tRNA relay from one ribosome face to the next, or with relaying tRNAs between successive ribosomes in a polysome. Additional experiments are required to validate this model and to determine if it also represents a general mechanism in higher eukaryotes.

It is unclear whether the two modes of polysome association reported for yeast vigilin are related. Although it is possible that Bfr1p serves as a bridge between vigilin and Asc1p/RACK1 at the ribosome, the requirement of both Bfr1p and Asc1p/RACK1 for vigilin's association with polysomes[45,64] argues in favor of the vigilin-Bfr1p-Pab1p complex being distinct from the vigilin-Asc1p/RACK1-eEF1A complex. In agreement, the presence of Pab1p (which interacts with the initiation factor eIF4G) in one complex and the elongation factor eEF1A in the other suggest that the complex act on different stages of translation.

Translation Repression by Vigilin

Surprisingly, vigilin has also been reported to repress translation. In human breast cancer, vigilin reduces metastasis in part by repressing translation of the proto-oncogene c-*fms*, although the mechanism(s) through which this occurs is unclear.[4] More is known on its role in translation repression in the budding yeast. Here, vigilin accomplishes this as part of a protein network named SESA, for its protein components: Smy2p/GIGYF2, Eap1p, Scp160p/vigilin, and Asc1p/RACK1.[24] Smy2p is a GYF domain protein identified as a high-copy suppressor of deletion mutants of *MPS2*, an integral membrane protein involved in insertion of the spindle pole body (the yeast analog of metazoan centrosomes) into the nuclear envelope.[24] Eap1p is a yeast family member of the eIF4E-binding proteins (4E-BPs) that generally function as inhibitors of translation initiation by competing with eIF4G for eIF4E-binding.[70] Eap1p and vigilin show genetic interaction, since loss of both genes in *S. cerevisiae* is lethal.[71] Interestingly, this synthetic lethality can only be rescued by a version of Eap1p competent for eIF4E-binding, suggesting that its functional relationship with vigilin is related to translation.[71] Moreover, the SESA network member Smy2p does not co-immunoprecipitate with an Eap1p mutant defective for eIF4E-binding.[24]

Upon defects in SPB duplication, the SESA network represses translation of *POM34* mRNA, which encodes a component of the nuclear pore complex (NPC).[24] As NPC assembly and SPB duplication share a common machinery,[72] it is hypothesized that the decreased amount of Pom34p avails more of the limited factors for SPB duplication.[24,73] Given the presence of Eap1 in the SESA network, translation inhibition of *POM34* mRNA likely occurs at the stage of initiation. In support, *POM34* mRNA is normally localized at the ER where its translation occurs and becomes mislocalized in yeast mutants of the initiation factor eIF4E.[24] Unfortunately, the mechanistic role of vigilin and Asc1p/RACK1 in translation repression of *POM34* has not been assessed further. Indeed, binding of *POM34* mRNA by the SESA network is mediated by Smy2p (and not vigilin), and whether the Eap1p-Scp160p/vigilin-Asc1p/RACK1 complex can repress other mRNAs is unclear. Interestingly, a homologous complex (vigilin-RACK1-GIGYF2) with translation repression capabilities was identified in human cells.[74] Notably, two vigilin-bound mRNAs—*DHH1* and *YOR338W*—whose polysome occupation was reduced in the absence of vigilin[75] did not co-immunoprecipitate with Smy2p.[24] This observation indicates the presence of at least two ways for recruiting mRNAs to the Eap1p-Scp160p/vigilin-Asc1p/RACK1 translation repression complex; one of which works via Smy2p as an adaptor, and the other via direct binding to vigilin.

Regulation of mRNA Stability by Vigilin

Aside from translation, vigilin can also act on the stability of its target mRNAs. The first example of this role was found in *X. laevis*. Here, the vitellogenin mRNA, which encodes an egg yolk precursor protein and homolog of apolipoproteins,[76] is stabilized in the male liver upon estrogen stimulation.[35,77] This stabilization is achieved by vigilin-binding to a 94 nt sequence in the 3′UTR of vitellogenin mRNA, which

also contains cleavage sites for the PMR-1 endonuclease.[26,36] *In vitro* cleavage assays show that vigilin stabilizes vitellogenin mRNA by blocking access of PMR-1 to the cleavage sites.[26] Moreover, as the cleavage assay was performed with recombinant proteins, vigilin is sufficient for stabilization of vitellogenin mRNA.

Competition of vigilin with another RBP for a binding site is also the basis for regulation of the proto-oncogene c-*fms* mRNA in human breast cancer.[4] Here, vigilin competes with HuR, a member of the ELAV family of RBPs,[4,78,79] but in contrast to vitellogenin mRNA, upregulated expression of vigilin induces the destabilization of c-*fms* mRNA. HuR has been reported to bind and stabilize a number of mRNAs[80] and the competition of the two proteins for the same 69 nt element in the c-*fms* 3′UTR could regulate c-*fms* mRNA translation and stability.[4] The similarities of these two examples suggest that vigilin's function in regulating mRNA stability is to restrict effector proteins' access to target mRNAs.

This kind of regulation seems unlikely to reflect a general function of vigilin since high-throughput analysis of mouse liver transcripts have shown no observable changes in the steady-state levels of the transcriptome upon knockdown or overexpression of vigilin.[2] However, vigilin also influences formation of P-bodies in *S. cerevisiae*.[28] P-bodies are stress-induced mRNPs containing factors of the translation repression and decay machineries, and thought to be sites for mRNA translation repression and degradation.[81] In yeast mutants lacking vigilin, P-body formation becomes dysregulated and occurs even under nonstress conditions.[28] Strikingly, these ectopic P-bodies are not properly formed and lack bona-fide P-body constituents.[28] Since nontranslating mRNA is a requirement of P-body formation,[82] it is possible that vigilin regulates P-body assembly in part by impacting the translation status of subsets of mRNAs, although this remains to be tested.

Bringing a and α Together: Vigilin in mRNA Localization

Mating in *S. cerevisiae* is triggered by reciprocal detection of cell-type-specific pheromones and begins with cell polarization, followed by directional growth of the triggered cells towards each other.[83] Early on, mRNAs encoding polarity and secretion factors (POLs) are transported to the site proximal to the pheromone, where their locally translated products function to establish a cell projection called a shmoo.[22] Localization of these mRNAs is dependent on both vigilin and Myo4p, a myosin motor.[22,84]

Myo4p is a well-characterized component of the RNA translocation machinery that localizes *ASH1* mRNA into the mature bud.[85] For *ASH1* mRNA, binding is facilitated by the RBPs She2p and She3p.[86] However, localization of POL mRNAs is dependent on the binding in their 3′UTRs by vigilin's KH14 domain, and truncation of KH14 abolishes both mRNA-binding and POL mRNA localization.[22] Vigilin might, therefore, replace She2p/She3p during transport of POL mRNAs, which is supported by its reported physical interaction with Myo4p.[22]

Interestingly, vigilin can also function as an effector of G-protein-coupled receptor (GPCR) signaling in yeast mating.[87] GPCRs are a large class of cell surface receptors activated by a number of different stimuli, including mating pheromones.[83,88] Upon stimulation, the yeast α subunit (Gpa1p) of the GPCR-associated trimeric G-protein exchanges its bound guanosine diphosphate (GDP) for guanosine triphosphate (GTP) and dissociates from the β/γ dimer subunit. Vigilin interacts with the activated GTP-bound form of Gpa1p, but not the inactive GDP-bound form.[87] Possibly, GTP-Gpa1p anchors vigilin at the shmoo tip, which helps to concentrate vigilin-bound POL mRNAs at this site and thus facilitates local translation of POL factors (Figure 3(a)). In agreement with its roles as a G-protein effector and in POL mRNA localization, vigilin deletion mutants fail to orient their shmoos toward pheromone sources and show severely reduced mating efficiency.[22,87]

A COMMON FUNCTION OF VIGILINS?

Despite a wealth of reports on vigilin functions, we lack a clear understanding of if and how these observations are mechanistically linked, or whether vigilins are promiscuous RBPs that change function in a context-dependent manner (Figure 3). The field is facing several challenges to be able to answer this question. One of these challenges lies in the promiscuity of nucleic acids bound by vigilins (Box 2). Even if the most recent consensus sequence identified by PAR-CLIP[2] represents the true vigilin RNA-binding sequence, its high degeneracy suggests that vigilins may be rather general RNA binders which function in context with and/or regulate association of other RBPs. This notion is supported by the variety of RNA species bound by vigilin (Box 2). As described, mRNA-binding has been shown for vigilins in *H. sapiens*,[38,47] *M. musculus*,[2] *X. laevis*,[26,35,36] *D. melanogaster*,[23,38] and *S. cerevisiae*.[1,22,25,75] tRNA-binding has also been reported for the human[34,54] and mouse[2] vigilins. In addition, the interaction of

the yeast vigilin with eEF1A[45]—which delivers aminoacyl-tRNAs to the ribosome A-site—provides indirect evidence that the yeast homolog may also associate with tRNAs. Although vigilin can bind different RNA species, it is unknown if a single vigilin molecule binds a homogeneous class of RNAs. Whereas vigilins in an RNA-transporting mRNP may bind a homogeneous class of mRNA (or tRNA), polysome-associated vigilins may bind both the translated mRNAs as well as tRNAs. The identification of RNAs bound by vigilins in different protein complexes could provide insight in this regard. However, vigilins' many KH domains and relatively short consensus target sequence adds an additional layer of complexity to their mode of RNA-binding. Do the KH domains merely contribute to overall affinity for the same RNA molecule which contains multiple copies of the target sequence or do they mediate scaffolding of the same or different transcript(s)? The RNA-binding data currently available do not allow us to differentiate between these possibilities. Understanding the mode of RNA-binding of vigilins across different protein complexes and different organisms will undoubtedly shed light on how vigilins' RNA-related functions are coordinated.

Another challenge is determining the individual contribution of each KH domain to vigilin's overall function. While many vigilin-dependent functions require the presence of the C-terminal KH10-14 domains, it remains unclear what roles, if any, the N-terminal domains play. Although the functions of vigilin's C-terminal KH domains have been well investigated, similarly in-depth analyses of the N-terminal KH domain is lacking. However, with both the NES and NLS located in vigilin's N-terminal half, targeted deletion or site-directed mutagenesis of N-terminal KH domains may be required to accurately characterize their contribution to vigilin function. The observation that an ectopic KH13-14 fragment acts as a gain-of-function mutant that disrupts chromosome segregation and heterochromatin formation[19] suggests that vigilin's N-terminus may confer a regulatory or balancing function via interaction with other targets. Obviously, the large number of KH domains raises the possibility of functional redundancy. A more detailed structure–function analysis would address this possibility and help to associate the different KH domains with various vigilin functions. Potentially, a pattern of how vigilin coordinates its functions among its domain architecture could emerge from such studies.

By extension, structure–function analyses across vigilin homologs, as well as substitution of homologous KH domains, would also address the evolution of vigilin's roles among different organisms. Given the varying assortment of classical to divergent KH domains between homologs of lower and higher eukaryotes (Figure 2), one interesting question would be whether the evolution of divergent KH domains to classical KH domains altered the proteins' mechanisms of function and/or conferred novel functions? Domain swapping or complementation assays like those performed by Cortés et al.[37] would address such questions and help to determine which of the vigilin's functions have been mechanistically conserved.

Our current knowledge of vigilin is disparate, owing to the many roles vigilin plays and complicated by the different functions between homologs. By addressing the challenges highlighted, we can begin to bring the field of vigilin research together toward a more comprehensive understanding how the proteins' functions are coordinated and how they have evolved.

REFERENCES

1. Hogan DJ, Riordan DP, Gerber AP, Herschlag D, Brown PO. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol* 2008, 6:e255.

2. Mobin MB, Gerstberger S, Teupser D, Campana B, Charisse K, Heim MH, Manoharan M, Tuschl T, Stoffel M. The RNA-binding protein vigilin regulates VLDL secretion through modulation of Apob mRNA translation. *Nat Commun* 2016, 7:12848.

3. Yang WL, Wei L, Huang WQ, Li R, Shen WY, Liu JY, Xu JM, Li B, Qin Y. Vigilin is overexpressed in hepatocellular carcinoma and is required for HCC cell proliferation and tumor growth. *Oncol Rep* 2014, 31:2328–2334.

4. Woo H-H, Yi X, Lamb T, Menzl I, Baker T, Shapiro DJ, Chambers S. Posttranscriptional suppression of proto-oncogene c-fms expression by vigilin in breast cancer. *Mol Cell Biol* 2011, 31:215–225.

5. Wintersberger U, Kühne C, Karwan A. Scp160p a new yeast protein associated with the nuclear membrane and the endoplasmic reticulum, is necessary for maintenance of exact ploidy. *Yeast* 1995, 11:929–944.

6. Goolsby KM, Shapiro DJ. RNAi-mediated depletion of the 15 KH domain protein, vigilin, induces death of dividing and non-dividing human cells but does not initially inhibit protein synthesis. *Nucleic Acids Res* 2003, 31:5644–5653.

7. Siomi H, Matunis MJ, Michael WM, Dreyfuss G. The pre-messenger RNA binding K-protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res* 1993, 21:1193–1198.

8. Valverde R, Edwards L, Regan L. Structure and function of KH domains. *FEBS J* 2008, 275:2712–2726.

9. Nicastro G, Taylor IA, Ramos A. KH-RNA interactions: back in the groove. *Curr Opin Struct Biol* 2015, 30C:63–70.

10. Tomizawa T, Kigawa T, Koshiba S, Inoue M, Yokoyama S. Solution structure of the 14th KH type I domain from human Vigilin. Available at: https://www.rcsb.org/pdb/explore/explore.do?structureId=2CTM.

11. Musco G, Stier G, Joseph C, Morelli MAC, Nilges M, Gibson TJ, Pastore A. NMR study of vigilin, Repeat 6, minimized average structure. Available at: https://www.rcsb.org/pdb/explore/explore.do?structureId=1VIH.

12. Tomizawa T, Kigawa T, Koshiba S, Inoue M, Yokoyama S. Solution structure of the 13th KH type I domain from human Vigilin. Available at: https://www.rcsb.org/pdb/explore/explore.do?structureId=2CTL.

13. Tomizawa T, Kigawa T, Koshiba S, Inoue M, Yokoyama S. Solution structure of the 12th KH type I domain from human Vigilin. Available at: https://www.rcsb.org/pdb/explore/explore.do?structureId=2CTK.

14. Tomizawa T, Kigawa T, Koshiba S, Inoue M, Yokoyama S. Solution structure of the 8th KH type I domain from human Vigilin. Available at: https://www.rcsb.org/pdb/explore/explore.do?structureId=2CTJ.

15. Tomizawa T, Kigawa T, Koshiba S, Inoue M, Yokoyama S. Solution structure of the 4th KH type I domain from human Vigilin. Available at: https://www.rcsb.org/pdb/explore/explore.do?structureId=2CTF.

16. Tomizawa T, Kigawa T, Koshiba S, Inoue M, Yokoyama S. Solution structure of the 1st KH type I domain from human Vigilin. Available at: https://www.rcsb.org/pdb/explore/explore.do?structureId=2CTE.

17. Weber V, Wernitznig A, Hager G, Harata M, Frank P, Wintersberger U. Purification and nucleic-acid-binding properties of a *Saccharomyces cerevisiae* protein involved in the control of ploidy. *Eur J Biochem* 1997, 249:309–317.

18. Huertas D, Cortés A, Casanova J, Azorín F. Drosophila DDP1, a multi-KH-domain protein, contributes to centromeric silencing and chromosome segregation. *Curr Biol* 2004, 14:1611–1620.

19. Zhou J, Wang Q, Chen L-L, Carmichael GG. On the mechanism of induction of heterochromatin by the RNA-binding protein vigilin. *RNA* 2008, 14:1773–1781.

20. Wei L, Xie X, Li J, Li R, Shen W, Duan S, Zhao R, Yang W, Liu Q, Fu Q, et al. Disruption of human vigilin impairs chromosome condensation and segregation. *Cell Biol Int* 2015, 9999:1–8.

21. Irie K, Tadauchi T, Takizawa PA, Vale RD, Matsumoto K, Herskowitz I. The Khd1 protein, which has three KH RNA-binding motifs, is required for proper localization of ASH1 mRNA in yeast. *EMBO J* 2002, 21:1158–1167.

22. Gelin-Licht R, Paliwal S, Conlon P, Levchenko A, Gerst JE. Scp160-dependent mRNA trafficking mediates pheromone gradient sensing and chemotropism in yeast. *Cell Rep* 2012, 1:483–494.

23. Nelson MR, Luo H, Vari HK, Cox BJ, Simmonds AJ, Krause HM, Lipshitz H, Smibert C. A multiprotein complex that mediates translational enhancement in *Drosophila*. *J Biol Chem* 2007, 282:34031–34038.

24. Sezen B, Seedorf M, Schiebel E. The SESA network links duplication of the yeast centrosome with the protein translation machinery. *Genes Dev* 2009, 23:1559–1570.

25. Hirschmann WD, Westendorf H, Mayer A, Cannarozzi G, Cramer P, Jansen R-P. Scp160p is required for translational efficiency of codon-optimized mRNAs in yeast. *Nucleic Acids Res* 2014, 42:4043–4055.

26. Cunningham KS, Dodson RE, Nagel MA, Shapiro DJ, Schoenberg DR. Vigilin binding selectively inhibits cleavage of the vitellogenin mRNA 3′-untranslated region by the mRNA endonuclease polysomal ribonuclease 1. *Proc Natl Acad Sci U S A* 2000, 97:12498–12502.

27. Wen W-L, Stevenson AL, Wang C-Y, Chen H-J, Kearsey SE, Norbury CJ, Watt S, Bahler J, Wang S-W. Vgl1, a multi-KH domain protein, is a novel component of the fission yeast stress granules required for cell survival under thermal stress. *Nucleic Acids Res* 2010, 38:6555–6566.

28. Weidner J, Wang C, Prescianotto-Baschong C, Estrada AF, Spang A. The polysome-associated proteins Scp160 and Bfr1 prevent P body formation under

normal growth conditions. *J Cell Sci* 2014, 127:1992–2004.

29. Graham DL, Oram JF. Identification and characterization of a high density lipoprotein-binding protein in cell membranes by ligand blotting. *J Biol Chem* 1987, 262:7439–7442.

30. McKnight GL, Reasoner J, Gilbert T, Sundquist KO, Hokland B, McKernan PA, Champagne J, Johnson CJ, Bailey MC, Holly R, et al. Cloning and expression of a cellular high density lipoprotein-binding protein that is up-regulated by cholesterol loading of cells. *J Biol Chem* 1992, 267:12131–12141.

31. Schmidt C, Henkel B, Pöschl E, Zorbas H, Purschke WG, Gloe TR, Müller PK. Complete cDNA sequence of chicken vigilin, a novel protein with amplified and evolutionary conserved domains. *Eur J Biochem* 1992, 206:625–634.

32. Neu-Yilik G, Zorbas H, Gloe TR, Raabe H-M, Hopp-Christensen TA, Müller PK. Vigilin is a cytoplasmic protein. A study on its expression in primary cells and in established cell lines of different species. *Eur J Biochem* 1993, 213:727–736.

33. Plenz G, Kügler S, Schnittger S, Rieder H, Fonatsch C, Müller PK. The human vigilin gene: identification, chromosomal localization and expression pattern. *Hum Genet* 1994, 93:575–582.

34. Kruse C, Grünweller A, Notbohm H, Kügler S, Purschke WG, Müller P. Evidence for a novel cytoplasmic tRNA-protein complex containing the KH-multidomain protein vigilin. *Biochem J* 1996, 320:247–252.

35. Dodson RE, Shapiro DJ. An estrogen-inducible protein binds specifically to a sequence in the 3′ untranslated region of estrogen-stabilized vitellogenin mRNA. *Mol Cell Biol* 1994, 14:3130–3138.

36. Dodson RE, Shapiro DJ. Vigilin, a ubiquitous protein with 14 K homology domains, is the estrogen-inducible vitellogenin mRNA 3′-untranslated region-binding protein. *J Biol Chem* 1997, 272:12249–12252.

37. Cortés A, Huertas D, Fanti L, Pimpinelli S, Marsellach FX, Piña B, Azorín F. DDP1, a single-stranded nucleic acid-binding protein of *Drosophila*, associates with pericentric heterochromatin and is functionally homologous to the yeast Scp160p, which is involved in the control of cell ploidy. *EMBO J* 1999, 18:3820–3833.

38. Wang Q, Zhang Z, Blackwell K, Carmichael GG. Vigilins bind to promiscuously A-to-I-edited RNAs and are involved in the formation of heterochromatin. *Curr Biol* 2005, 15:384–391.

39. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 2008, 36: W465–W469.

40. Kruse C, Emmrich J, Rumpel E, Klinger MH, Grünweller A, Rohwedel J, Kramer H-J, Kühnel W, Müller PK. Production of trypsin by cells of the exocrine pancreas is paralleled by the expression of the KH protein vigilin. *Exp Cell Res* 1998, 239:111–118.

41. Vollbrandt T, Willkomm D, Stossberg H, Kruse C. Vigilin is co-localized with 80S ribosomes and binds to the ribosomal complex through its C-terminal domain. *Int J Biochem Cell Biol* 2004, 36:1306–1318.

42. Currie JR, Brown WT. KH domain-containing proteins of yeast: absence of a fragile X gene homologue. *Am J Med Genet* 1999, 84:272–276.

43. Brykailo MA, McLane LM, Fridovich-Keil J, Corbett AH. Analysis of a predicted nuclear localization signal: implications for the intracellular localization and function of the *Saccharomyces cerevisiae* RNA-binding protein Scp160. *Nucleic Acids Res* 2007, 35:6862–6869.

44. Kramer K, Sachsenberg T, Beckmann BM, Qamar S, Boon K-L, Hentze MW, Kohlbacher O, Urlaub H. Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. *Nat Methods* 2014, 11:1064–1070.

45. Baum S, Bittins M, Frey S, Seedorf M. Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes. *Biochem J* 2004, 380:823–830.

46. Li A-M, Vargas CA, Brykailo MA, Openo KK, Corbett AH, Fridovich-Keil JL. Both KH and non-KH domain sequences are required for polyribosome association of Scp160p in yeast. *Nucleic Acids Res* 2004, 32:4768–4775.

47. Kanamori H, Dodson RE, Shapiro DJ. *In vitro* genetic analysis of the RNA binding site of vigilin, a multi-KH-domain protein. *Mol Cell Biol* 1998, 18:3991–4003.

48. Jette N, Lees-Miller SP. The DNA-dependent protein kinase: a multifunctional protein kinase with roles in DNA double strand break repair and mitosis. *Prog Biophys Mol Biol* 2015, 117:194–205. 10.1016/j.pbiomolbio.2014.12.003.

49. Kügler S, Grünweller A, Probst C, Klinger M, Müller P, Kruse C. Vigilin contains a functional nuclear localisation sequence and is present in both the cytoplasm and the nucleus. *FEBS Lett* 1996, 382:330–334.

50. Klinger MH, Kruse C. Immunocytochemical localization of vigilin, a tRNA-binding protein, after cell fractionation and within the exocrine pancreatic cell of the rat. *Ann Anat* 1996, 178:331–335. 10.1016/S0940-9602(96)80086-0.

51. Frey S, Pool M, Seedorf M. Scp160p, an RNA-binding, polysome-associated protein, localizes to the endoplasmic reticulum of *Saccharomyces cerevisiae* in a

microtubule-dependent manner. *J Biol Chem* 2001, 276:15905–15912.

52. Batlle M, Marsellach F-X, Huertas D, Azorín F. Drosophila vigilin, DDP1, localises to the cytoplasm and associates to the rough endoplasmic reticulum. *Biochim Biophys Acta* 2011, 1809:46–55.

53. Kruse C, Willkomm DK, Grünweller A, Vollbrandt T, Sommer S, Busch S, Pfeiffer T, Brinkmann J, Hartmann RK, Müller PK. Export and transport of tRNA are coupled to a multi-protein complex. *Biochem J* 2000, 346:107–115.

54. Kruse C, Grünweller A, Willkomm DK, Pfeiffer T, Hartmann RK, Müller PK. tRNA is entrapped in similar, but distinct, nuclear and cytoplasmic ribonucleoprotein complexes, both of which contain vigilin and elongation factor 1 alpha. *Biochem J* 1998, 329:615–621.

55. Marsellach F-X, Huertas D, Azorín F. The multi-KH domain protein of *Saccharomyces cerevisiae* Scp160p contributes to the regulation of telomeric silencing. *J Biol Chem* 2006, 281:18227–18235.

56. Piacentini L, Fanti L, Negri R, Del Vescovo V, Fatica A, Altieri F, Pimpinelli S. Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in *Drosophila*. *PLoS Genet* 2009, 5:e1000670.

57. Zeng W, Ball AR, Yokomori K. HP1: heterochromatin binding proteins working the genome. *Epigenetics* 2010, 5:287–292.

58. Saksouk N, Simboeck E, Déjardin J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* 2015, 8:3.

59. Becker JS, Nicetto D, Zaret KS. H3K9me3-dependent heterochromatin: barrier to cell fate changes. *Trends Genet* 2016, 32:29–41. 10.1016/j.tig.2015.11.001.

60. Muramatsu D, Kimura H, Kotoshiba K, Tachibana M, Shinkai Y. Pericentric H3K9me3 formation by HP1 interaction-defective histone methyltransferase Suv39h1. *Cell Struct Funct* 2016, 41:145–152.

61. Hickman MA, Froyd CA, Rusche LN. Reinventing heterochromatin in budding yeasts: Sir2 and the origin recognition complex take center stage. *Eukaryot Cell* 2011, 10:1183–1192.

62. Liu Q, Yang B, Xie X, Wei L, Liu W, Yang W, Ge Y, Zhu Q, Zhang J, Jiang L, et al. Vigilin interacts with CCCTC-binding factor (CTCF) and is involved in CTCF-dependent regulation of the imprinted genes Igf2 and H19. *FEBS J* 2014, 281:2713–2725.

63. Lewis A, Murrell A. Genomic imprinting: CTCF protects the boundaries. *Curr Biol* 2004, 14:R284–R286.

64. Lang BD, Li A-M, Black-Brewster HD, Fridovich-Keil JL. The brefeldin A resistance protein Bfr1p is a component of polyribosome-associated mRNP

65. Lang BD, Fridovich-Keil JL. Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. *Nucleic Acids Res* 2000, 28:1576–1584.

66. Hollingworth D, Candel AM, Nicastro G, Martin SR, Briata P, Gherzi R, Ramos A. KH domains with impaired nucleic acid binding as a tool for functional analysis. *Nucleic Acids Res* 2012, 40:6873–6886.

67. Cannarozzi G, Schraudolph NN, Faty M, von Rohr P, Friberg MT, Roth AC, Gonnet P, Gonnet G, Barral Y. A role for codon order in translation dynamics. *Cell* 2010, 141:355–367. 10.1016/j.cell.2010.02.036.

68. Negrutskii BS, Deutscher MP. Channeling of aminoacyl-tRNA for protein synthesis *in vivo*. *Proc Natl Acad Sci U S A* 1991, 88:4991–4995.

69. Gavin A-C, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dümpelfeld B, et al. Proteome survey reveals modularity of the yeast cell machinery. *Nature* 2006, 440:631–636.

70. Cosentino GP, Schmelzle T, Haghighat A, Helliwell SB, Hall MN, Sonenberg N. Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in *Saccharomyces cerevisiae*. *Mol Cell Biol* 2000, 20:4604–4613.

71. Mendelsohn BA, Li A-M, Vargas CA, Riehman K, Watson A, Fridovich-Keil JL. Genetic and biochemical interactions between SCP160 and EAP1 in yeast. *Nucleic Acids Res* 2003, 31:5838–5847.

72. Casey AK, Dawson TR, Chen J, Friederichs JM, Jaspersen SL, Wente SR. Integrity and function of the *Saccharomyces cerevisiae* spindle pole body depends on connections between the membrane proteins Ndc1, rtn1, and yop1. *Genetics* 2012, 192:441–455.

73. Witkin KL, Friederichs JM, Cohen-Fix O, Jaspersen SL. Changes in the nuclear envelope environment affect spindle pole body duplication in *Saccharomyces cerevisiae*. *Genetics* 2010, 186:867–883.

74. Belozerov VE, Ratkovic S, McNeill H, Hilliker AJ, McDermott JC. *In vivo* interaction proteomics reveal a novel p38 mitogen-activated protein kinase/Rack1 pathway regulating proteostasis in *Drosophila* muscle. *Mol Cell Biol* 2014, 34:474–484.

75. Li A-M, Watson A, Fridovich-Keil JL. Scp160p associates with specific mRNAs in yeast. *Nucleic Acids Res* 2003, 31:1830–1837.

76. Baker ME. Is vitellogenin an ancestor of apolipoprotein B-100 of human low-density lipoprotein and human lipoprotein lipase? *Biochem J* 1988, 255:1057–1060.

77. Shelness GS, Sellers JA. Very-low-density lipoprotein assembly and secretion. *Curr Opin Lipidol* 2001, 12:151–157.

78. Woo H-H, Zhou Y, Yi X, David C, Zheng W, Gilmore-Hebert M, Kluger HM, Ulukus EC, Baker T,

Stoffer JB, et al. Regulation of non-AU-rich element containing c-fms proto-oncogene expression by HuR in breast cancer. *Oncogene* 2009, 28:1176–1186.

79. Woo H-H, Chambers KS. Post-transcriptional regulation of proto-oncogene c-fms in breast cancer. In: Siregar Y, *Oncogene and Cancer – From Bench to Clinic*. Rijeka, Croatia: InTech; 2013.

80. Hinman M, Lou H. Diverse molecular functions of Hu proteins. *Cell Mol Life Sci* 2008, 65:3168–3181.

81. Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. *Mol Cell* 2007, 25:635–646.

82. Teixeira D, Sheth U, Valencia-Sanchez MA, Brengues M, Parker R. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* 2005, 11:371–382.

83. Merlini L, Dudin O, Martin SG. Mate and fuse: how yeast cells do it. *Open Biol* 2013, 3:130008.

84. Aronov S, Gelin-Licht R, Zipor G, Haim L, Safran E, Gerst JE. mRNAs encoding polarity and exocytosis factors are cotransported with the cortical endoplasmic reticulum to the incipient bud in *Saccharomyces cerevisiae*. *Mol Cell Biol* 2007, 27:3441–3455.

85. Böhl F, Kruse C, Frank A, Ferring D, Jansen R-P. She2p, a novel RNA-binding protein tethers ASH1 mRNA to the Myo4p myosin motor via She3p. *EMBO J* 2000, 19:5514–5524.

86. Edelmann FT, Schlundt A, Heym RG, Jenner A, Niedner-Boblenz A, Syed MI, Paillart J-C, Stehle R, Janowski R, Sattler M, et al. Molecular architecture and dynamics of ASH1 mRNA recognition by its mRNA-transport complex. *Nat Struct Mol Biol* 2017, 24:152–161.

87. Guo M, Aston C, Burchett SA, Dyke C, Fields S, Rajarao SJR, Uetz P, Wang Y, Young K, Dohlman HG. The yeast G protein α subunit Gpa1 transmits a signal through an RNA binding effector protein Scp160. *Mol Cell* 2003, 12:517–524.

88. Hanlon CD, Andrew DJ. Outside-in signaling—a brief review of GPCR signaling with a focus on the *Drosophila* GPCR family. *J Cell Sci* 2015, 128:3533–3542.

9.2 – Research Article – The RNA-Binding Protein Scp160p Facilitates Aggregation of Many Endogenous Q/N-Rich Proteins. (Cheng et al., 2018)

The RNA-Binding Protein Scp160p Facilitates Aggregation of Many Endogenous Q/N-Rich Proteins

Graphical Abstract
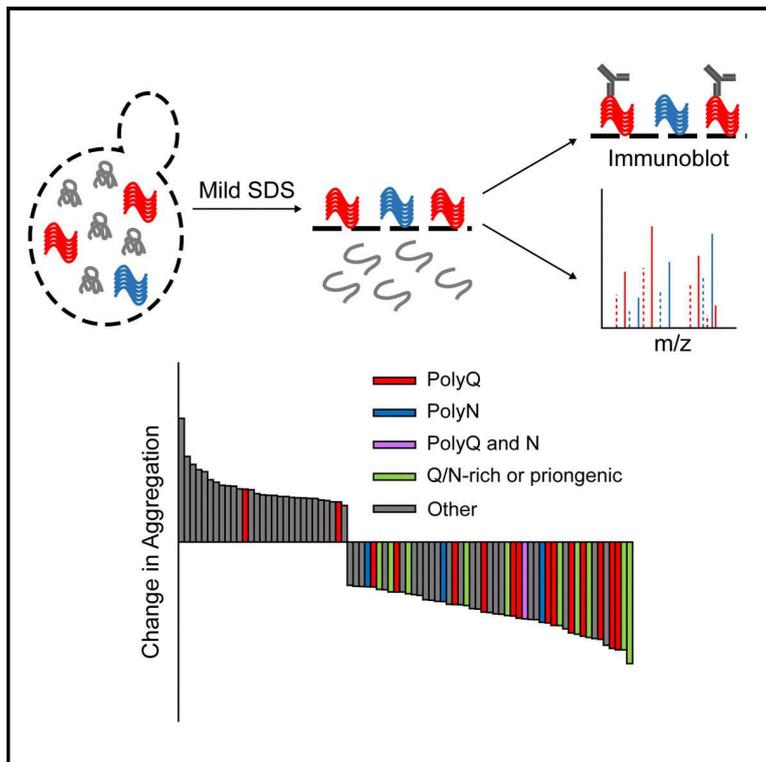
Authors

Matthew H.K. Cheng,
Patrick C. Hoffmann,
Mirita Franz-Wachtel, Carola Sparn,
Charlotte Seng, Boris Maček,
Ralf-Peter Jansen

Correspondence

ralf.jansen@uni-tuebingen.de

In Brief

Glutamine and asparagine (Q/N)-mediated protein aggregation can lead to neurodegenerative diseases but is also functionally important. Cheng et al. report a method to assess aggregation of proteins at the proteomic level and observe reduced aggregation of many endogenous Q/N-rich proteins in the absence of the yeast vigilin protein.

Highlights

- Yeast RNA-binding protein Scp160p facilitates aggregation of polyQ reporters

- Codon usage has a major influence in polyQ-induced toxicity

- A combined method to assess protein aggregation at the proteome level

- Loss of Scp160p results in reduced aggregation of many Q/N-rich proteins

Data and Software Availability

PXD008175

The RNA-Binding Protein Scp160p Facilitates Aggregation of Many Endogenous Q/N-Rich Proteins

Matthew H.K. Cheng,[1,2] Patrick C. Hoffmann,[2,3] Mirita Franz-Wachtel,[4] Carola Sparn,[2,5] Charlotte Seng,[2] Boris Maček,[1,4] and Ralf-Peter Jansen[1,2,6,*]

[1]International Max Planck Research School "From Molecules to Organisms," Spemannstraße 35, 72076 Tübingen, Germany
[2]Interfaculty Institute of Biochemistry, University of Tübingen, Hoppe-Seyler Straße 4, 72076 Tübingen, Germany
[3]MRC Laboratory of Molecular Biology, Francis Crick Avenue, CB2 0QH Cambridge, UK
[4]Proteome Center Tübingen, University of Tübingen, Auf der Morgenstelle 15, 72076 Tübingen, Germany
[5]DKFZ, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
[6]Lead Contact
*Correspondence: ralf.jansen@uni-tuebingen.de
https://doi.org/10.1016/j.celrep.2018.06.015

SUMMARY

The RNA-binding protein Scp160p is the yeast homolog of the conserved vigilin protein family. These proteins influence a variety of nuclear and cytoplasmic functions. One of Scp160p's reported roles is to increase translation elongation efficiency in a manner related to codon usage. Thus, it can affect translation speed and co-translational folding of nascent peptides. We used polyglutamine (polyQ) reporters to assess Scp160p's effect on protein synthesis and observed that, in the absence of Scp160p, aggregation of polyQ is reduced and toxicity is abolished. We additionally took a proteomic approach and analyzed the impact of Scp160p on the aggregation of endogenous proteins under normal growth conditions. In the absence of Scp160p, aggregation of many Q/N-rich proteins was reduced. Because aggregation mediated by these regions can be important for the proteins' functions, Scp160p may affect many processes via aggregation of Q/N-rich proteins.

INTRODUCTION

The RNA-binding protein (RBP) Scp160p is the yeast homolog of the conserved vigilin family of proteins (Cheng and Jansen, 2017; McKnight et al., 1992). These RBPs are characterized by their domain architecture, comprising 14–15 nucleic acid-binding heterogeneous nuclear ribonucleoprotein particle (hnRNP) K homology (KH) domains (Siomi et al., 1993). With hundreds of RNA targets (Hogan et al., 2008; Mobin et al., 2016), these proteins influence many cellular processes (Cheng and Jansen, 2017). In *S. cerevisiae*, Scp160p associates with polysomes via the small ribosomal subunit constituent Asc1p (RACK1 in mammals) (Baum et al., 2004) and has been linked to ploidy maintenance, mating, and mRNA translation, transport, and decay (Cheng and Jansen, 2017). It has also been implicated in toxicity associated with expanded polyglutamine (polyQ) expression (Kaiser et al., 2013). We previously reported a role for Scp160p

in enhancing elongation efficiency in the context of mRNA codon usage (Hirschmann et al., 2014).

By influencing elongation kinetics, codon usage can regulate protein synthesis by varying elongation speed. This promotes proper folding of local peptide regions, ensures natively folded and functional proteins (Buhr et al., 2016; Yu et al., 2015), and is guided by both synonymous codon selection and codon order along the coding region (Cannarozzi et al., 2010). Scp160p is proposed to hinder diffusion of deacylated-tRNAs from translating polysomes, allowing their efficient recharge and reuse (Hirschmann et al., 2014). Although substrate channeling acts on tRNAs whose synonymous cognate codons are ordered in succession (autocorrelated), Scp160p acts primarily on tRNAs whose synonymous cognate codons alternate and are sporadic (nonautocorrelated) (Hirschmann et al., 2014). To study Scp160p's impact on protein synthesis pertaining to codon usage, we assessed how loss of Scp160p affects the biology of polyQ regions, which can be encoded by two codons: CAG and CAA.

Expanded polyQ and glutamine/asparagine (Q/N)-rich regions have been linked to a variety of neurodegenerative diseases. Their pathogenicity is believed to result from their roles in the aggregation of their encompassing proteins. These aggregates can sequester chaperones and overwhelm the protein quality control system, in addition to disturbing membrane integrity (Park et al., 2013; Sakahira et al., 2002). Despite their pathogenicity, Q- and Q/N-rich regions are prevalent in the eukaryotic proteome (Michelitsch and Weissman, 2000; Schaefer et al., 2012). In their native nonexpanded context, polyQ regions and their aggregation are functionally important (Fiumara et al., 2010; Gemayel et al., 2015; Khan et al., 2015; Lee et al., 2013; O'Rourke and Reines, 2016), and Q/N richness is a common feature of low-complexity regions (LCRs) that drives protein phase separation (Bergeron-Sandoval et al., 2016).

Here, we show that both aggregation and toxicity of three polyQ reporters are influenced by Scp160p. We also assessed the impact of Scp160p on the aggregation of endogenous proteins in the yeast proteome and show that its loss affects the aggregation of many endogenous Q/N-rich proteins, which might explain the diverse phenotypes seen in *scp160Δ* cells.
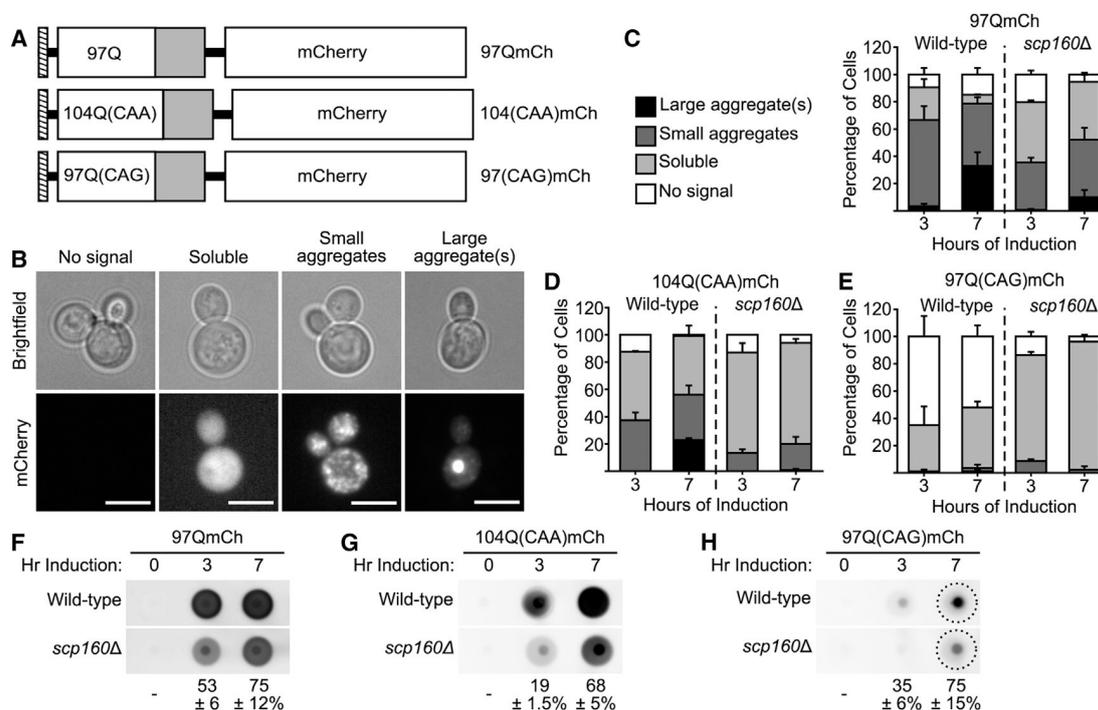
Figure 1. Reduced Accumulation of PolyQ Aggregates in *scp160Δ* Cells

(A) Three polyQ reporters used in this study. All three carry an N-terminal myc tag (hatch), the Htt proline-rich region (gray), and a C-terminal mCherry tag (labeled).

(B) Cells were assigned to one of four classes based on mCherry signal distribution. Scale bar, 5 μm.

(C–E) Percentage of cells in each of the four classes after induction of 97QmCh (C), 104Q(CAA)mCh (D), and 97Q(CAG)mCh (E). Error bars show SEM of three biological replicates.

(F–H) Filter trap binding shows that aggregation of 97QmCh (F), 104Q(CAA)mCh (G), and 97Q(CAG)mCh (H) is reduced in *scp160Δ* cells. Numbers show the immunosignal from *scp160Δ* cells as a fraction (±SEM of three biological replicates) of that from wild-type cells. Immunosignals were measured from regions of interest with consistent sizes (dashed circle in H).

RESULTS

Accumulation of PolyQ Aggregates Is Reduced in *scp160Δ* Mutants Regardless of Codon Usage

To investigate Scp160p's role in protein synthesis, we employed three reporters derived from exon 1 of expanded huntingtin protein (mHtt)—97QmCh, 104Q(CAA)mCh, and 97Q(CAG)mCh—encoding the same amino acids but differing in codon usage of the polyQ region. The 104Q(CAA)mCh and 97Q(CAG)mCh reporters, whose polyQ regions are encoded solely by the frequent CAA or the infrequent CAG codons, are highly autocorrelated. In contrast, the polyQ region of 97QmCh, encoded by both CAA and CAG codons, is nonautocorrelated. The galactose-inducible reporters are tagged at the N and C termini with myc and mCherry, respectively (Figure 1A).

97QmCh, 104Q(CAA)mCh, or 97Q(CAG)mCh was induced in wild-type and *scp160Δ* cells, and aggregation was assessed by microscopy after 3 and 7 hr. The reporters existed in three discernible states: soluble, multiple small aggregates, or one to two large aggregates (Figure 1B). Some cells lacked a detectable mCherry signal after 7 hr of induction. The percentage of wild-type and *scp160Δ* cells in each of these states approximated the reporters' level of aggregation.

97QmCh and 104Q(CAA)mCh aggregated readily in wild-type cells but to a lesser extent in *scp160Δ* cells. After 7 hr of induction, 97QmCh localized into foci in most wild-type cells but only half of *scp160Δ* cells (Figure 1C). Similarly, for 104Q(CAA)mCh, approximately half the wild-type cells contained aggregates after 7 hr of induction, whereas only a fifth of *scp160Δ* cells did (Figure 1D). In contrast, 97Q(CAG)mCh expression was poor in wild-type cells, with no detectable signal in half of the cells after 7 hr of induction (Figure 1E). Western blots against the N-terminal myc tag of 97Q(CAG)mCh show that the lack of signal is not due to loss of the mCherry tag from frameshifting (Figure S1). Expression of 97Q(CAG)mCh is more robust in *scp160Δ* cells, with a soluble mCherry signal detected in nearly all observed cells (Figure 1E).

One phenotype associated with *scp160Δ* cells is increased ploidy. To rule out that reduced polyQ aggregation results from this phenotype, we assessed 97QmCh aggregation in cells from which Scp160p was depleted by a tetracycline-regulated system (Tet-off *SCP160*) but before ploidy increase (Figure S2). After 7 hr of induction in Tet-off *SCP160* cells, distribution of 97QmCh was similar to that observed in *scp160Δ* cells (Figure S2), suggesting reduced polyQ aggregation is independent of ploidy phenotype.

Figure 2. mHtt Toxicity Depends on Codon Usage and a Start Codon

(A) 97QmCh and 104Q(CAA)mCh are modestly toxic to both wild-type and *scp160Δ* cells. 97Q(CAG)mCh is highly toxic to wild-type cells, but not *scp160Δ* cells.

(B) Toxicity of 97Q(CAG)mCh is lost after removing all in-frame start codons upstream of the polyCAG.

We also employed filter trap binding to assess aggregation of the polyQ reporters (Wanker et al., 1999). This assay exploits the SDS-resistant nature of polyQ aggregates to trap them on nitrocellulose membrane.

Consistent with our microscopy, filter trap binding showed that 97QmCh and 104Q(CAA)mCh aggregated in wild-type cells and that their aggregation was reduced in *scp160Δ* cells (Figures 1F and 1G). Moreover, we observed by filter trap binding that 97Q(CAG)mCh aggregation is reduced in *scp160Δ* cells compared to wild-type (Figure 1H). Aggregation of all three reporters was not abolished in *scp160Δ* cells but was similarly reduced. Thus, a factor besides Scp160p likely contributes to polyQ aggregation.

Altogether, our imaging and biochemical data show that polyQ-mediated aggregation depends on Scp160p function. The extra 7 Qs of 104Q(CAA)mCh did not alter this effect in *scp160Δ* cells (Figure S3). Because the polyQ regions of these reporters differ in codon usage and autocorrelation, Scp160p's role in polyQ-mediated aggregation is independent of that in codon-related translation.

RNA Sequence and Translation of the PolyQ Region Affects Toxicity in Yeast

Because loss of Scp160p reduced aggregation of the three polyQ reporters, we asked whether it may also influence toxicity. Induction of 97Q(CAG)mCh conferred severe toxicity to wild-type cells, in contrast to 97QmCh and 104Q(CAA)mCh (Figure 2A). Because the protein sequences of the reporters are the same, the RNA sequence—specifically of the polyQ region—affects toxicity. 97Q(CAG)mCh is the only accurate representation of mHtt at both the nucleic acid and the protein levels. This RNA species is proposed to self-anneal into a long stem that can be bound unspecifically by RBPs, sequestering them from normal functions (de Mezer et al., 2011). Alternatively, polyCAG repeats have been reported to trigger repeat-associated non-ATG (RAN) translation, resulting in toxic peptides from spontaneous initiation in all six reading frames within the repeat

(Bañez-Coronel et al., 2015). To test whether ATG-initiated translation is required for mHtt toxicity, we generated (-)ATG-97Q(CAG)mCh, in which the two ATG codons upstream of the polyCAG were mutated to stop codons. Because these mutations are outside the polyCAG, the stem formed by the repeat should remain unaffected (Figure S1C). Toxicity is lost from (-)ATG-97Q(CAG)mCh in wild-type cells (Figure 2B), arguing that an ATG upstream of the polyCAG repeat is required for toxicity. Similar to 97Q(CAG)mCh, expression of (-)ATG-97Q(CAG)mCh did not affect growth of *scp160Δ* cells (Figure 2B). Therefore, in our experimental system, polyQ-induced toxicity does not primarily result from RNA secondary structures.

Scp160p Facilitates Aggregation of Many Endogenous Q/N-Rich Proteins

We next asked whether Scp160p also affects aggregation of endogenous polyQ-containing proteins under normal growth. To this end, we combined filter trap binding with quantitative mass spectrometry to assess changes in the aggregation state of wild-type and *scp160Δ* proteomes. Aggregated proteins were isolated by filter trap binding and subsequently digested on the nitrocellulose membrane with trypsin. The resulting peptides were recovered, labeled with dimethyl light or dimethyl heavy, and analyzed by nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS) (Figure 3A; Experimental Procedures). This allowed comparison of the relative abundance of proteins from wild-type and *scp160Δ* cells that remain trapped on the membrane and thus their aggregation states. The comparison is displayed as a $\log_2$ ratio of the signal intensity from *scp160Δ* cells over wild-type ($\log_2(scp160Δ/\text{wild-type})$). In parallel, we subjected the input whole-cell lysates (WCLs) to the same MS analysis, which allowed us to normalize observed intensities to changes in the protein's abundance between wild-type and *scp160Δ* cells. The unnormalized and normalized $\log_2(scp160Δ/\text{wild-type})_{\text{filter-trapped}}$ ratios showed a Pearson correlation = 0.79 (Figure S4A), suggesting that observed changes in the aggregation state of proteins in *scp160Δ* cells are not due to altered protein abundance. Nevertheless, subsequent analyses were performed with normalized $\log_2(scp160Δ/\text{wild-type})$ ratios.

Our analysis revealed that the aggregation state of 77 proteins was significantly ($p_{\text{sigB}} < 0.1$) altered in *scp160Δ* compared to wild-type cells (Figure 3B). 27 of the 48 (56.3%) proteins with reduced aggregation in *scp160Δ* cells contained a Q/N-rich
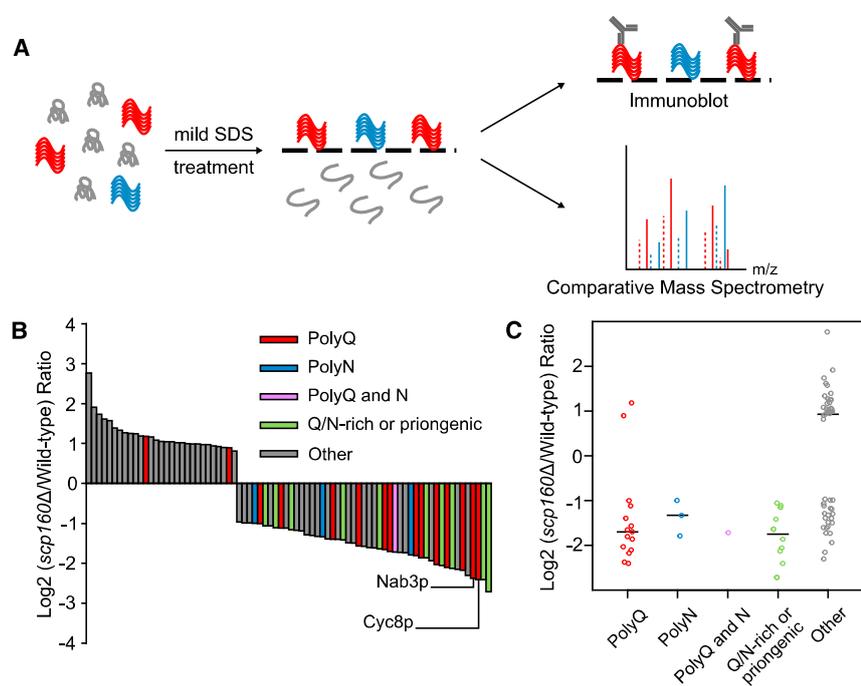
Figure 3. Reduced Aggregation of Many Endogenous Q/N-Rich Proteins in scp160Δ

(A) Workflow of the filter trap binding and MS method.

(B) Aggregation of 77 proteins was altered in scp160Δ cells compared to wild-type cells. Those with reduced aggregation in scp160Δ cells are enriched for Q/N-rich proteins. Cyc8p and Nab3p were selected for validation by conventional filter trap binding and microscopy. See also Table S1.

(C) Distribution of the $\log_2(scp160\Delta/\text{wild-type})_{\text{filter-trapped}}$ ratios for the Q/N-rich proteins. Bars show the median.

Cyc8p and Nab3p were genomically tagged at their C termini with either GFP or 3xmyc. Filter trap binding showed reduced Cyc8-GFP aggregation in scp160Δ cells compared to wild-type cells (Figure 4B). This effect was also observed by semi-denaturing detergent agarose electrophoresis (Figure S5A). Normalization of the filter trap binding to western blot of the whole-cell lysate demonstrated that reduced Cyc8-GFP aggregation (63% ± 5.9% of the wild-type signal) is not due to a change in the protein's abundance. Similarly, for Nab3p, aggregation of Nab3-3xmyc in scp160Δ cells was reduced to 45% ± 6.0% of that in wild-type cells (Figure 4C). These results are in full agreement with our proteomic analysis. Genomic PCRs of the polyQ-encoding regions in CYC8 and NAB3 show that this effect is not caused by altered polyQ lengths (Figure S5B).

As an orthogonal method to assess Cyc8p and Nab3p aggregation, we compared the distribution of Cyc8-GFP and Nab3-GFP in wild-type and scp160Δ cells. We calculated a ratio of the maximum to minimum signals (signal intensity ratio) of the area inside the cell containing GFP (Figure 4D). A signal intensity ratio = 1 would indicate uniform GFP signal distribution, while a higher signal intensity ratio would indicate the presence of at least one denser or more concentrated GFP focus. The distribution of signal intensity ratios between wild-type and scp160Δ cells for Cyc8-GFP revealed a statistically significant change between the two populations (p = 0.007978, n = 81; Mann-Whitney U test). This is consistent with our filter trap binding. A change in the distribution of signal intensity ratios for Nab3-GFP could also be observed in scp160Δ cells, but it was not statistically significant (p = 0.06572, n = 85; Mann-Whitney U test). Altogether, our observations by filter trap binding and microscopy support our proteomic data and demonstrate that under normal growth conditions, Scp160p promotes the aggregation of the polyQ-containing proteins Cyc8p and Nab3p.

DISCUSSION

One aspect by which protein synthesis can be regulated is codon usage, which can influence elongation kinetics and thus protein folding and function (Buhr et al., 2016; Yu et al., 2015).

region or were predicted to form prions (Figures 3B and 3C; Table S1). This represents a 30-fold enrichment over the predicted number of all yeast proteins containing Q/N-rich regions (1.69%) (Michelitsch and Weissman, 2000) and suggests that Scp160p has a widespread role in facilitating aggregation of Q/N-rich proteins.

A look at the molecular weights and isoelectric points (pIs) of proteins whose aggregation changed upon loss of Scp160p showed that our approach did not bias toward these parameters (Figure S4B). The Q/N-rich proteins that are affected by loss of Scp160p span several cellular functions, including transcription, RNA processing and translation, nucleocytoplasmic transport, vesicle trafficking, and phosphorylation (Table S1; Michelitsch and Weissman, 2000; Saccharomyces Genome Database [SGD]). The diverse functions associated with these Q/N-rich proteins thus argue against bias of the method or scp160Δ toward specific pathways or functions.

Loss of Scp160p Leads to Reduced Aggregation of the PolyQ-Containing Proteins Cyc8p and Nab3p

We validated our proteomic data with two polyQ-containing proteins by conventional filter trap binding and microscopy. Cyc8p and Nab3p were chosen because they were among the Q/N-rich proteins whose aggregation was most reduced in scp160Δ cells (Figure 3B) and because they hold different molecular functions (Table S1). Cyc8p is a component of a general transcription-regulating complex (Tanaka and Mukai, 2015), while Nab3p is an RBP and subunit of a complex involved in transcription termination and 3′ processing of several classes of RNAs (Loya et al., 2013). The polyQ regions of both Cyc8p and Nab3p are important for their biological functions (Gemayel et al., 2015; O'Rourke and Reines, 2016).

A Cyc8p

Nab3p

B Wild-type scp160Δ

Cyc8p — FT

Pgk1p — WB

63
± 5.9%

C Wild-type scp160Δ

Nab3p — FT

Pgk1p — WB

45
± 6.0%

D

GFP (Fire)   Hoechst   Brightfield

Wild-type

scp160Δ

Cyc8-GFP

Wild-type

scp160Δ

Nab3-GFP

E

Cyc8-GFP

Signal Intensity Ratio (Max/Min)

Wild-type   scp160Δ

p = 0.007978

Nab3-GFP

Signal Intensity Ratio (Max/Min)

Wild-type   scp160Δ

p = 0.06572

Figure 4. Aggregation of Cyc8p and Nab3p Is Reduced in scp160Δ

(A) Schematic of Cyc8p and Nab3p. Glutamines are shaded in red, and asparagines are in blue.

(B and C) Filter trap binding shows reduced aggregation of Cyc8-GFP (B) and Nab3-3xmyc (C) in scp160Δ cells. Western blots (WBs) show little change in the proteins' abundance in wild-type and scp160Δ cells. Numbers show the normalized immunosignal from scp160Δ cells as a fraction (±SEM of three biological replicates) of that from wild-type cells.

(D) Distribution of GFP signals displayed with "Fire" pseudo-coloring. Enlarged images of the nuclei indicated by arrowheads are shown on left.

(E) Distribution of signal intensity ratios for Cyc8-GFP (n = 81) and Nab3-GFP (n = 85) in wild-type and scp160Δ cells.

Our analyses demonstrate that polyQ-induced toxicity, but not protein aggregation, depends on codon usage. A requirement for the presence of an ATG codon upstream of the polyCAG repeat argues that toxicity of 97Q(CAG)mCh in our system is unlikely to result from toxic products of RAN translation or RNA secondary structures. This requirement of translation for toxicity, combined with its separation from aggregation, may hint that differences in elongation—as influenced by codon usage—can lead the three polyQ reporters to adopt distinct folds and consequent conformations that differ in toxicity. Studies have shown that conformation, and not merely aggregation, is a major determinant for polyQ-induced toxicity (Miller et al., 2011; Shen et al., 2016). Aggregation of the polyQ reporters is decreased in the absence of Scp160p possibly due to slowed translation (Hirschmann et al., 2014). Scp160p's role in facilitating polyQ aggregation appears to be separate from its role in autocorrelation.

Given the prevalence of polyQ regions and LCRs like Q/N-rich stretches in the eukaryotic proteome (Schaefer et al., 2012), we combined filter trap binding and quantitative MS to assess whether Scp160p could also influence the aggregation states of endogenous Q/N-rich proteins. This assay should be applicable to all cell types and provides an additional tool to rapidly and globally compare SDS-resistant protein aggregation in response to treatments or changes in genetic background.

Reduced aggregation upon loss of Scp160p was especially prevalent for Q/N-rich proteins, among them the prionogenic proteins Sup35p ([PSI+]), New1p ([NU+]), and Cyc8p ([OCT+]) (Glover et al., 1997; Liebman and Chernoff, 2012; Patel et al., 2009). Scp160p has already been implicated in [PSI+] (Sup35p) prion induction, although the effect was not reproduced in a re-engineered deletion (Manogaran et al., 2011).

During our validation of Cyc8p and Nab3p aggregation, we realized that microscopy provided a good measure to compare the density of foci, although it could not account for other attributes, such as size, number, and compactness. Therefore, an analysis by imaging leads to underrepresentation of changes in Cyc8p and Nab3p aggregation. The notable difference in the amplitude of protein aggregation in scp160Δ versus wild-type cells, as measured by imaging and filter trap binding, might stem from the higher sensitivity of the filter trap binding in detecting aggregation. Nevertheless, the analyses of Cyc8p and Nab3p using biochemical and imaging approaches together provide strong evidence that the aggregation of these proteins is changed in the absence of Scp160p.

The biological importance of polyQ and Q/N-rich regions is increasingly apparent for many proteins (Gemayel et al., 2015; Khan et al., 2015; Lee et al., 2013; O'Rourke and Reines, 2016). The repeat length of Cyc8p's middle Q-rich region (Figure 4A) modulates not only the protein's aggregation propensity but also the expression of target genes and consequently flocculation (Gemayel et al., 2015). Although this effect pertains to different lengths in the Q-rich regions between both natural and engineered variants, regulation of polyQ-mediated aggregation could represent an alternative means of control. Strikingly, among the 27 reported genes whose expression are regulated by Cyc8p's polyQ length (Gemayel et al., 2015), 7 were previously observed to change in expression levels upon depletion of Scp160p (Schreck, 2010). Thus, while polyQ regions can

contribute to transcriptional control, Scp160p could act as an additional layer to fine-tune transcription factor activity and target gene expression. By extension, Scp160p may have a similar impact on Nab3p function. Although the biological role of Nab3p's polyQ region is unclear, it has been shown to self-assemble into amyloid polymers and hydrogels *in vitro*. The polyQ region's importance is underscored by the inviability of cells in which Nab3p aggregation has been disrupted by reducing the number of Qs in this region (O'Rourke and Reines, 2016).

Q/N-rich LCRs are present in proteins that span a range of functions, including transcription, RNA metabolism, nucleocyto-plasmic transport, vesicle trafficking, and phosphorylation (Michelitsch and Weissman, 2000; Schaefer et al., 2012). This raises the possibility for many cellular processes to be regulated rapidly and reversibly via protein aggregation. Our data demonstrate a role for Scp160p in promoting Q/N-mediated protein aggregation under normal conditions, which may thus affect their normal functions. Such a scenario would account for the many seemingly unrelated phenotypes associated with *scp160Δ* cells. Understanding the mechanistic function of Scp160p on Q/N-rich proteins will provide additional insight into the nuance of protein aggregation and its impact on cell function.

EXPERIMENTAL PROCEDURES

Filter Trap Binding

10–40 μg (immunoblotting) or 400 μg (proteomic studies) of whole-cell lysates were semi-denatured by treatment with PBS + 1% SDS at room temperature for 2 hr (Wanker et al., 1999). Treated lysates were passed through a nitrocel-lulose membrane (Whatman Protran, pore size 0.45 μm) using a Minifold I vacuum filtration system (Schleicher & Schuell). The membrane was rinsed twice with PBS + 1% SDS and twice with PBS. Immunoblotting was done per standard methods.

For proteomic studies, proteins were digested on the membrane with trypsin as described previously (Luque-Garcia and Neubert, 2009). The areas where lysates were applied were cut out, rinsed four times with water, and washed for 5 min at room temperature in water with agitation. Membranes were blocked in polyvinylpyrrolidone with an average molecular weight of 40 (PVP-40) (Sigma-Aldrich) for 30 min at 37°C with gentle agitation. PVP-40 was removed, and the membranes were washed six times with water. To digest the proteins, 250 ng of trypsin in 20 μL of 50 mM ammonium bicarbonate was added. Additional ammonium bicarbonate was added to submerge the membranes. Digestion proceeded overnight at 37°C with gentle agitation.

Samples were dried by speed vacuum. 90 μL of acetone per 4 mm^2 of membrane was added to each sample, vortexed, and incubated 30 min at room temperature. Peptide pellets were precipitated at $14,000 \times g$ for 10 min, air-dried, and then resuspended in 20 μL of sample solution (2% ACN and 0.1% formic acid).

Nano-LC-MS/MS Analysis and MS Data Processing

Samples were labeled with dimethyl light ($(CH_3)_2$) and dimethyl heavy ($(CH_1D_2)_2$) reagents and measured on an Easy-nLC 1200 System (Thermo Fisher Scientific) coupled to a Q Exactive HF mass spectrometer (Thermo Fisher Scientific), as described elsewhere (Kliza et al., 2017). For details on sample preparation, MS measurement, and data processing, refer to Supple-mental Information.

Perseus software (v.1.5.0.15), a module from the MaxQuant suite (Tyanova et al., 2016), was used to calculate the significance B (p_{sigB}) for each protein ratio with respect to the distance from the median of the distribution of all protein ratios, as well as intensities. All proteins with $p_{sigB} < 0.1$ in a pairwise comparison were considered differentially retained on the nitrocellulose membrane.

Q/N richness was determined by inspection of the protein sequences. PolyQ regions were defined as a stretch with at least 6 glutamines for every 8 consec-utive amino acids (8-mers, allowing 2 mismatches) (Totzeck et al., 2017). The same parameter was used to define polyN regions. Q/N-rich and prionogenic proteins were those computationally predicted and published in at least Michelitsch and Weissman (2000) or Alberti et al. (2009).

Statistical Analysis

Distributions of signal intensity ratios were compared by Mann-Whitney U test (wilcox.test function) in R.

DATA AND SOFTWARE AVAILABILITY

The accession number for the proteomic data reported in this paper is PRIDE: PXD008175.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and one table and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.06.015.

AUTHOR CONTRIBUTIONS

M.H.K.C. and R.-P.J. conceived the project. M.H.K.C. performed the experiments, analyzed the data, and wrote the manuscript. C. Sparn, P.C.H., and C. Seng performed the experiments. M.F.-W. and B.M. designed, performed, and analyzed the MS experiments. R.P.-J. supervised the project, interpreted the data, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of priono-genic proteins. Cell *137*, 146–158.

Bañez-Coronel, M., Ayhan, F., Tarabochia, A.D., Zu, T., Perez, B.A., Tusi, S.K., Pletnikova, O., Borchelt, D.R., Ross, C.A., Margolis, R.L., et al. (2015). RAN Translation in Huntington Disease. Neuron *88*, 667–677.

Baum, S., Bittins, M., Frey, S., and Seedorf, M. (2004). Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes. Biochem. J. *380*, 823–830.

Bergeron-Sandoval, L.-P., Safaee, N., and Michnick, S.W. (2016). Mechanisms and Consequences of Macromolecular Phase Separation. Cell *165*, 1067–1079.

Buhr, F., Jha, S., Thommen, M., Mittelstaet, J., Kutz, F., Schwalbe, H., Rodnina, M.V., and Komar, A.A. (2016). Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations. Mol. Cell *61*, 341–351.

Cannarozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G., and Barral, Y. (2010). A role for codon order in translation dynamics. Cell *141*, 355–367.

Cheng, M.H., and Jansen, R.-P. (2017). A jack of all trades: the RNA-binding protein vigilin. Wiley Interdiscip. Rev. RNA 8, e1448.

de Mezer, M., Wojciechowska, M., Napierala, M., Sobczak, K., and Krzyzosiak, W.J. (2011). Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference. Nucleic Acids Res. 39, 3852–3863.

Fiumara, F., Fioriti, L., Kandel, E.R., and Hendrickson, W.A. (2010). Essential role of coiled coils for aggregation and activity of Q/N-rich prions and PolyQ proteins. Cell 143, 1121–1135.

Gemayel, R., Chavali, S., Pougach, K., Legendre, M., Zhu, B., Boeynaems, S., van der Zande, E., Gevaert, K., Rousseau, F., Schymkowitz, J., et al. (2015). Variable Glutamine-Rich Repeats Modulate Transcription Factor Activity. Mol. Cell 59, 615–627.

Glover, J.R., Kowal, A.S., Schirmer, E.C., Patino, M.M., Liu, J.-J., and Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of S. cerevisiae. Cell 89, 811–819.

Hirschmann, W.D., Westendorf, H., Mayer, A., Cannarozzi, G., Cramer, P., and Jansen, R.-P. (2014). Scp160p is required for translational efficiency of codon-optimized mRNAs in yeast. Nucleic Acids Res. 42, 4043–4055.

Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D., and Brown, P.O. (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol. 6, e255.

Kaiser, C.J.O., Grötzinger, S.W., Eckl, J.M., Papsdorf, K., Jordan, S., and Richter, K. (2013). A network of genes connects polyglutamine toxicity to ploidy control in yeast. Nat. Commun. 4, 1571.

Khan, M.R., Li, L., Pérez-Sánchez, C., Saraf, A., Florens, L., Slaughter, B.D., Unruh, J.R., and Si, K. (2015). Amyloidogenic Oligomerization Transforms Drosophila Orb2 from a Translation Repressor to an Activator. Cell 163, 1468–1483.

Kliza, K., Taumer, C., Pinzuti, I., Franz-Wachtel, M., Kunzelmann, S., Stieglitz, B., Macek, B., and Husnjak, K. (2017). Internally tagged ubiquitin: a tool to identify linear polyubiquitin-modified proteins by mass spectrometry. Nat. Methods 14, 504–512.

Lee, C., Zhang, H., Baker, A.E., Occhipinti, P., Borsuk, M.E., and Gladfelter, A.S. (2013). Protein aggregation behavior regulates cyclin transcript localization and cell-cycle control. Dev. Cell 25, 572–584.

Liebman, S.W., and Chernoff, Y.O. (2012). Prions in yeast. Genetics 191, 1041–1072.

Loya, T.J., O'Rourke, T.W., Degtyareva, N., and Reines, D. (2013). A network of interdependent molecular interactions describes a higher order Nrd1-Nab3 complex involved in yeast transcription termination. J. Biol. Chem. 288, 34158–34167.

Luque-Garcia, J.L., and Neubert, T.A. (2009). On-membrane tryptic digestion of proteins for mass spectrometry analysis. Methods Mol. Biol. 536, 331–341.

Manogaran, A.L., Hong, J.Y., Hufana, J., Tyedmers, J., Lindquist, S., and Liebman, S.W. (2011). Prion formation and polyglutamine aggregation are controlled by two classes of genes. PLoS Genet. 7, e1001386.

McKnight, G.L., Reasoner, J., Gilbert, T., Sundquist, K.O., Hokland, B., McKernan, P.A., Champagne, J., Johnson, C.J., Bailey, M.C., Holly, R., et al. (1992). Cloning and expression of a cellular high density lipoprotein-binding protein that is up-regulated by cholesterol loading of cells. J. Biol. Chem. 267, 12131–12141.

Michelitsch, M.D., and Weissman, J.S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. Proc. Natl. Acad. Sci. USA 97, 11910–11915.

Miller, J., Arrasate, M., Brooks, E., Libeu, C.P., Legleiter, J., Hatters, D., Curtis, J., Cheung, K., Krishnan, P., Mitra, S., et al. (2011). Identifying polyglutamine protein species in situ that best predict neurodegeneration. Nat. Chem. Biol. 7, 925–934.

Mobin, M.B., Gerstberger, S., Teupser, D., Campana, B., Charisse, K., Heim, M.H., Manoharan, M., Tuschl, T., and Stoffel, M. (2016). The RNA-binding protein vigilin regulates VLDL secretion through modulation of Apob mRNA translation. Nat. Commun. 7, 12848.

O'Rourke, T.W., and Reines, D. (2016). Determinants of Amyloid Formation for the Yeast Termination Factor Nab3. PLoS ONE 11, e0150865.

Park, S.-H., Kukushkin, Y., Gupta, R., Chen, T., Konagai, A., Hipp, M.S., Hayer-Hartl, M., and Hartl, F.U. (2013). PolyQ proteins interfere with nuclear degradation of cytosolic proteins by sequestering the Sis1p chaperone. Cell 154, 134–145.

Patel, B.K., Gavin-Smyth, J., and Liebman, S.W. (2009). The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion. Nat. Cell Biol. 11, 344–349.

Sakahira, H., Breuer, P., Hayer-Hartl, M.K., and Hartl, F.U. (2002). Molecular chaperones as modulators of polyglutamine protein aggregation and toxicity. Proc. Natl. Acad. Sci. USA 99 (Suppl 4), 16412–16418.

Schaefer, M.H., Wanker, E.E., and Andrade-Navarro, M.A. (2012). Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks. Nucleic Acids Res. 40, 4273–4287.

Schreck, H. (2010). Translational Control by the Multi KH Domain Protein Scp160 (Ludwig-Maximilians-Universität München).

Shen, K., Calamini, B., Fauerbach, J.A., Ma, B., Shahmoradian, S.H., Serrano Lachapel, I.L., Chiu, W., Lo, D.C., and Frydman, J. (2016). Control of the structural landscape and neuronal proteotoxicity of mutant Huntingtin by domains flanking the polyQ tract. eLife 5, 1–29.

Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucleic Acids Res. 21, 1193–1198.

Tanaka, N., and Mukai, Y. (2015). Yeast Cyc8p and Tup1p proteins function as coactivators for transcription of Stp1/2p-dependent amino acid transporter genes. Biochem. Biophys. Res. Commun. 468, 32–38.

Totzeck, F., Andrade-Navarro, M.A., and Mier, P. (2017). The Protein Structure Context of PolyQ Regions. PLoS ONE 12, e0170801.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods 13, 731–740.

Wanker, E.E., Scherzinger, E., Heiser, V., Sittler, A., Eickhoff, H., and Lehrach, H. (1999). Membrane filter assay for detection of amyloid-like polyglutamine-containing protein aggregates. Methods Enzymol. 309, 375–386.

Yu, C.-H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M.S., and Liu, Y. (2015). Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. Mol. Cell 59, 744–754.

Supplemental Information

The RNA-Binding Protein Scp160p Facilitates

Aggregation of Many Endogenous Q/N-Rich Proteins

Matthew H.K. Cheng, Patrick C. Hoffmann, Mirita Franz-Wachtel, Carola Sparn, Charlotte Seng, Boris Maček, and Ralf-Peter Jansen

A

Wild-type   scp160Δ   Wild-type   scp160Δ

Hr Induction 0 3 7 0 3 7   0 3 7 0 3 7

kDa
- 170
- 130
* - 100
- 70
▷ - 55

α-myc   α-mCherry

B Hr Induction   0   3   7

Wild-type scp160Δ Wild-type scp160Δ Wild-type scp160Δ

kDa
170 —
130 —
100 —   α-myc
70 —
55 —
40 —

α-Pgk1p

C

97Q(CAG)mCh
ΔG = -182.10 kcal/mol

(-)ATG-97Q(CAG)mCh
ΔG = -179.30 kcal/mol

97Q(CAG)mCh
GGA TCC **ATG** GAA CAA AAA CTT ATT TCT GAA GAA GAC TTG GGT GCT GGA
AGA TCT **ATG** GCG ACC CTG GAA AAG CTG **ATG** AAG GCC TTC GAG TCC CTC
AAA AGC TTC CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG ...

(-)ATG-97Q(CAG)mCh
GGA TCC GAA CAA AAA CTT ATT TCT GAA GAA GAC TTG GGT GCT GGA AGA
TCT **TAG** GCG ACC CTG GAA AAG CTG **TAG** AAG GCC TTC GAG TCC CTC AAA
AGC TTC CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG ...

BamHI - myc tag - N17 - polyQ

Figure S1 - 97Q(CAG)mCh does not undergo frameshifting and its secondary structure is not disrupted by mutation of in-frame start codons. Related to Figures 1 and 2.
A) Western blot against the N-terminal myc tag (left panel) and C-terminal mCherry tag (right panel) of 97Q(CAG)mCh show sharp distinct bands which argue that translation of the polyQ regions did not undergo frameshifting. The band at approximately 90kDa (*) represents the full length 97Q(CAG)mCh while the band at approximately 60kDa (arrowhead) represents 97Q(CAG)mCh reporter without its mCherry tag. **B)** After 7 hours induction, 97Q(CAG)mCh is more abundant in *scp160Δ* as compared to wild-type cells. **C)** RNAfold (Gruber et al., 2008) prediction show that mutating in-frame ATG start codons to TAG stop codons in (-)ATG-97Q(CAG)mCh does not disrupt the stem formed by the polyCAG repeat. Brackets indicate ATG start codons and arrowhead indicates start of the polyCAG repeat. The first ATG start codon (bolded black) of 97Q(CAG)mCh was deleted while the next two in-frame ATG codons upstream of the polyCAG repeats (bolded green) were mutated to TAG stop codons.
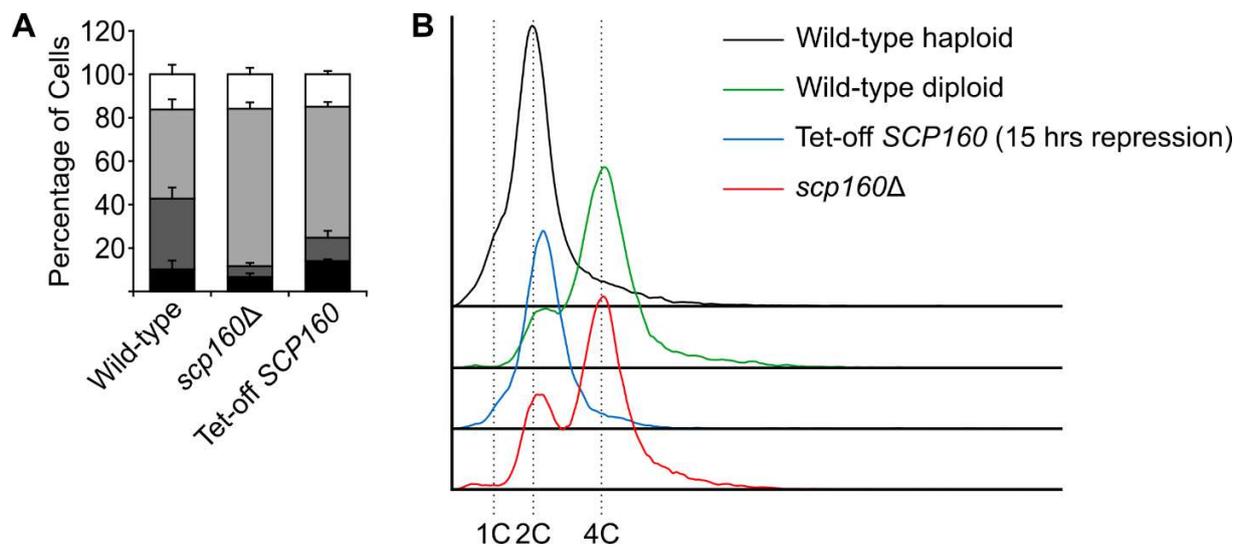
Figure S2 - Reduced 97QmCh aggregation is not due to ploidy phenotype of *scp160Δ*. Related to Figure 1.

SCP160 expression was shut off using a Tet-off system for 6 hours before induction of 97QmCh. 97QmCh was induced for 7 hours during which *SCP160* repression was maintained. **A)** 97QmCh aggregation is reduced when Scp160p is depleted. **B)** FACS analysis of DNA content using propidium iodide staining show that Tet-off *SCP160* cells were haploid when they were observed for 97QmCh aggregation. The 2C peaks of the wild-type haploid and Tet-off *SCP160* FACS profiles show that these cells were undergoing continuous growth.
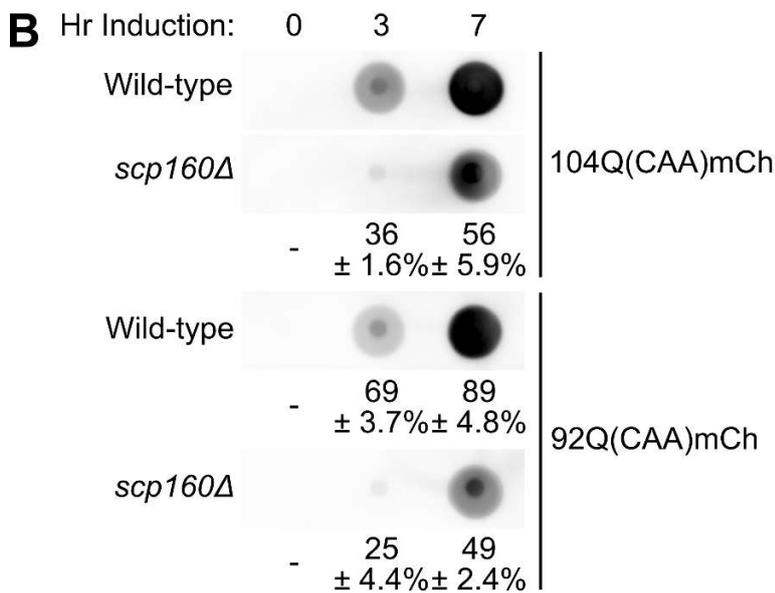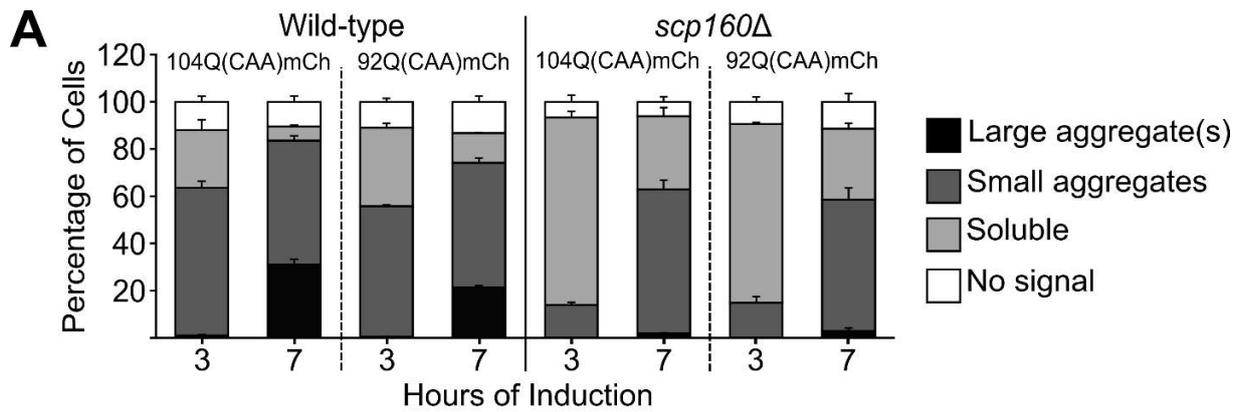
**Figure S3 - Additional 7 CAA's has negligible effect on aggregation in wild-type and *scp160Δ* cells. Related to Figure 1.**
A) The percentage of wild-type and *scp160Δ* cells in each of the four classes based on mCherry signal is negligible between 104Q(CAA)mCh and 92Q(CAA). B) Filter trap binding show the reductions in 104Q(CAA)mCh and 92Q(CAA)mCh in *scp160Δ* cells compared to wild-type are comparable.
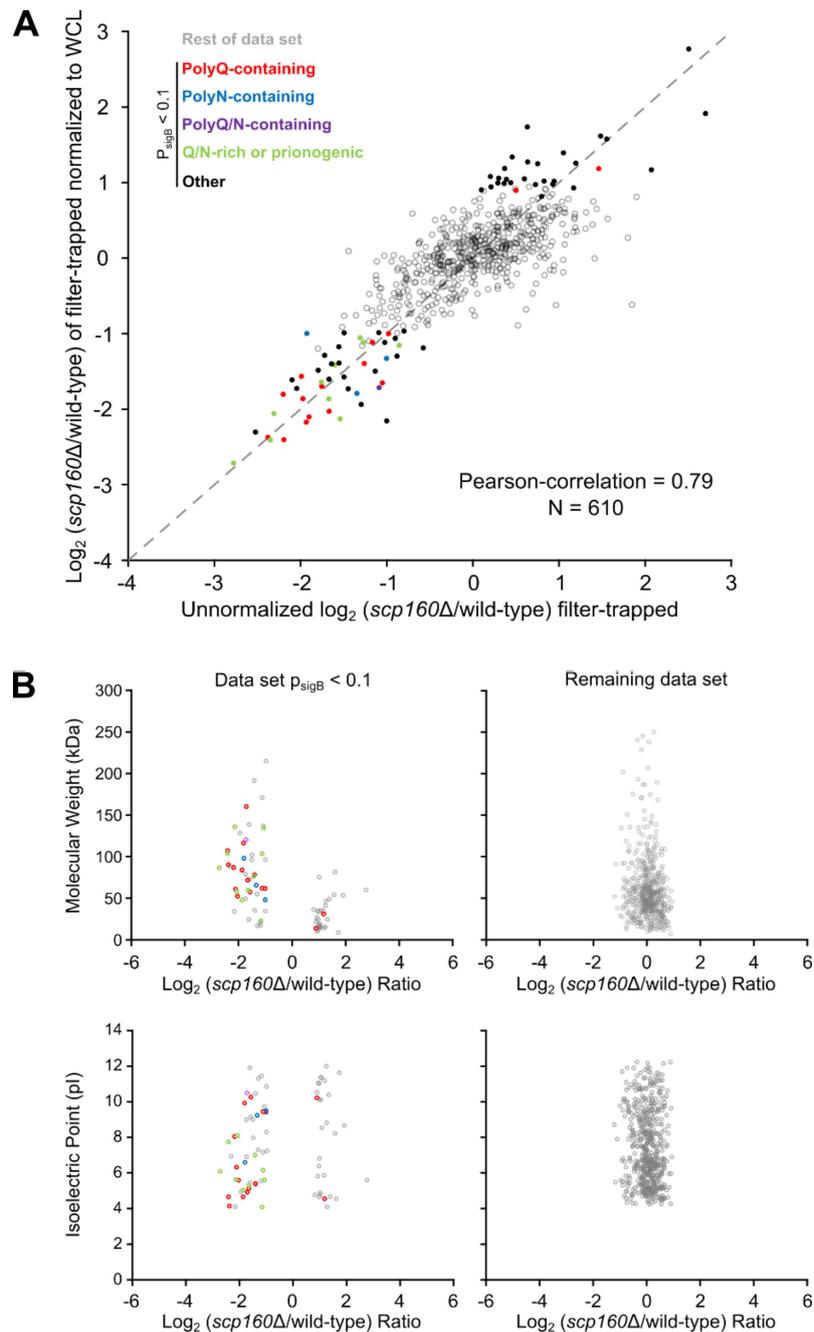
Figure S4 - Change in aggregation state is largely independent of different protein abundance and was not biased against protein size or charge. Related to Figure 3.
A) Comparison of the $\log_2(scp160\Delta/\text{wild-type})_{\text{filter-trapped}}$ ratios for 610 proteins before and after normalization with the corresponding $\log_2(scp160\Delta/\text{wild-type})_{\text{WCL}}$ ratios show a Pearson-correlation=0.79. This suggests that changes in aggregation state observed in $scp160\Delta$ cells is in general not due to different abundance of the proteins. **B)** Comparison of the proteins' $\log_2(scp160\Delta/\text{wild-type})$ ratios versus their molecular weight or isoelectric points show no correlation. Graphs on the left show proteins whose aggregation changed in $scp160\Delta$ cells with a $p_{\text{sigB}}<0.1$ while graphs on the right show all other proteins in our data set with $p_{\text{sigB}}>=0.1$.

Figure S5 - Reduced aggregation of Cyc8p and Nab3p are not due to changes in polyQ lengths. Related to Figure 4.
A) Semi-denaturing detergent agarose gel electrophoresis of Cyc8p-GFP show reduced amount of the SDS-resistant species in *scp160Δ* lysates compared to wild-type lysates. **B)** Genomic PCRs amplifying the genomic region encoding the polyQ sequences of Cyc8p and Nab3p show the lengths between wild-type and *scp160Δ* cells are unchanged.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Yeast strains

Yeast strains used in this study were of the W303 genetic background (*MAT*a *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1*). Strains were generated by PCR-based gene deletion (Longtine et al., 1998), genomic tagging (Knop et al., 1999), or one-step transformation with plasmids (Chen et al., 1992).

S. cerevisiae **strains used in this study**

Name	Genotype	Reference
W303a	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	-
RPY497	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1*	(Hirschmann et al., 2014)
RPY4584	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-97QmCh::URA3*	This study
RPY4588	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97QmCh::URA3*	This study
RPY4749	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-97Q(CAG)mCh::URA3*	This study
RPY4750	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-97Q(CAG)mCh::URA3*	This study
RPY4754	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-104Q(CAA)mCh::URA3*	This study
RPY4755	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-104Q(CAA)mCh::URA3*	This study
RPY4814	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::HIS3MX6 pTet-off-SCP160::klTRP1 pYES-97QmCh::URA3*	This study
RPY4867	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 CYC8-GFP::HIS3MX6*	This study
RPY4874	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 CYC8-GFP::HIS3MX6*	This study

RPY4993	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 NAB3-GFP::HIS3MX6*	This study
RPY5000	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 NAB3-GFP::HIS3MX6*	This study
RPY5072	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 NAB3-3xmyc::HIS3MX6*	This study
RPY5131	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-(-)ATG-97Q(CAG)mCh::URA3*	This study
RPY5132	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-(-)ATG-97Q(CAG)mCh::URA3*	This study
RPY5180	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::KANMX6 NAB3-3xmyc::HIS3MX6*	This study
RPY5271	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 pYES-92Q(CAA)mCh::URA3*	This study
RPY5272	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 scp160::klTRP1 pYES-92Q(CAA)mCh::URA3*	This study

Plasmid Construction

The pYES2-97QmCh plasmid was kindly provided by F Ulrich Hartl and was named p96QmCh in Park et al., 2013(Park et al., 2013). We named it 97QmCh for this study after sequencing revealed an additional glutamine in the polyQ region. pYES2-97QmCh as the template to generate the 97Q(CAG)mCh and 104Q(CAA)mCh variants. To create the 97Q(CAG)mCh and 104(CAA)mCh plasmids, the BglII to EcoRI fragment encoding the 97Q's were excised from pYES2-97QmCh and replaced with either a BglII to EcoRI 97CAG or BglII to EcoRI 104CAA fragment. The 97CAG and 104CAA fragments were synthesized by GeneArt Gene Synthesis service (Thermo Fisher Scientific). The (-)ATG-97Q(CAG)mCh variant was generated by replacing a BamHI to HindIII fragment with a BamHI to BglII myc-epitope encoding fragment in which the ATG codon was deleted, and a BglII to HindIII fragment encoding the N-terminal

17 amino acids of mHtt exon 1 in which all in frame ATG codons were mutated to TAG stop codons (Figure S3A).

Plasmids used in this study

Name	Description	Reference
Tet-off-*SCP160*	pCM182-*SCP160*	(Hirschmann et al., 2014)
97QmCh	pYES2-97QmCh::URA3	(Park et al., 2013)
97Q(CAG)mCh	pYES2-97Q(CAG)mCh::URA3	This study
92Q(CAA)mCh	pYES2-92Q(CAA)mCh::URA3	This study
104Q(CAA)mCh	pYES2-104Q(CAA)mCh::URA3	This study
(-)ATG-97Q(CAG)mCh	pYES2-(-)ATG-97Q(CAG)mCh::URA3	This study

PolyQ induction

PolyQ expression was induced from a pYES2 plasmid under a *GAL1* promoter. Cells were grown overnight at 30°C in selection media + 2% glucose to saturation and used to inoculate 80-100mL of selective media + 1% raffinose. The cultures recovered for 1 hour, allowing exhaustion of glucose from the pre-culture. Galactose was added at time 0 to 2%. At indicated times, 10-40 OD$_{600}$ units of cells (~1-4 x 10^8 cells) were collected for lysis.

Confocal microscopy

To prepare cells for microscopy, approximately 1 x 10^7 cells were fixed in 3.7% formaldehyde at room temperature for 15 min, then washed with 1x PBS. Cells were resuspended in 50-100μL of 1x PBS and stored at 4°C until use. For counterstaining of DNA, fixed cells were incubated for 30 min with 1x PBS containing 1μg/mL Hoechst stain (bisBenzimide H 33258; Sigma-Aldrich) in the dark, then washed twice with 1x PBS before microscopy. Cells were imaged on a Zeiss Axio Examiner.Z1 equipped with a CSU-X1 real-time confocal system, a VS-Laser system, and

a SPOT Flex CCD camera (Visitron Systems). Image acquisitions were made using VisiView software, with a 63x oil objective. Z-stacks had 30 steps of 0.3μm.

Analysis of Cyc8-GFP and Nab3-GFP distribution were performed in ImageJ. "Maximum Intensity" type Z-projections of the GFP Z-stacks were made. To obtain signal intensity values for GFP, a boundary was generated by thresholding areas containing GFP signal. The maximum and minimum signal intensities were measured using the "Analyze Particles" function. The ratio of maximum/minimum signal intensity was calculated to show the signal difference between foci and background staining, and as an approximation for aggregation.

Standard protein methods

To prepare yeast lysates, cells were collected at indicated time points by centrifugation for 5 min at 3000xg and 4°C. Cell pellets were washed 1x with 25mL of sterile water and flash frozen in liquid nitrogen and stored at -80°C until lysis. Lysis was carried on ice or at 4°C. Cell pellets were resuspended in 15μL/OD unit of lysis buffer (25mM Tris pH 7.5, 50mM KCl, 10mM MgCl2, 1mM EDTA, 5% glycerol, 0.5% triton X-100, 1x protease inhibitors (Roche cOmplete, EDTA-free tablets)), then lysed with glass beads on a VXR basic Vibrax shaker (IKA Works) for 6 cycles of 2 min with 1 min rest on ice in between. Lysates were cleared for 2 min at 1,200xg and 4°C.

For Western blots, whole cell lysates were boiled in Laemmli buffer for 5 min and then loaded onto an 8% SDS-polyacrylamide gel with a 6% stacking gel for analyses of mHtt reporter expression, and onto a 10% SDS-polyacrylamide gel with a 4% stacking gel for analyses of endogenous proteins.

Membranes were probed using mouse anti-c-myc 9E10 monoclonal (1:1,000 in 3% milk, Roche), mouse anti-mCherry monoclonal (1:1,000 in 3% milk, Clontech Living Colours), or

mouse anti-Pgk1p (1:10,000 in 3% milk, Invitrogen). HRP-conjugated sheep anti-mouse or HRP-conjugated goat anti-rabbit secondary antibodies were used. Detection was performed using ECL (Thermo Fisher Scientific) using an LAS-3000 Imager (Fujifilm), Fusion FX7 Chemiluminescence Imaging System (Vilber Lourmat), or ChemiDoc MP Imaging System (Bio-Rad). Signal quantification was done in ImageJ.

Semi-Denaturing Detergent Agarose Gel Electrophoresis

Semi-denaturing detergent agarose gel electrophoresis was performed as described by (Halfmann and Lindquist, 2008). Lysates were treated with sample buffer for 10 min at room temperature or at 95ºC to bring the proteins into monomeric form. Samples were run in a 1.6% agarose gel (dissolved in 1x TAE and supplemented with 1% SDS) at 100V for ~2 hours at 4 ºC. Transfer onto nitrocellulose (Whatman Protran, pore size 0.45µm) was done by capillary transfer overnight with 1x TBS.

NanoLC-MS/MS analysis and MS data processing

Extracted peptides were first desalted and then labeled using C18 StageTips (Boersema et al., 2013) as described elsewhere (Rappsilber et al., 2007). Samples were labeled with dimethyl "light" (($CH_3)_2$) and dimethyl "heavy" (($CH_1D_2)_2$), respectively. Complete incorporation levels of the dimethyl labels was achieved in all cases.

Eluted peptides were mixed in a 1:1 ratio according to measured protein amounts. The analysis of the peptide mixture was performed on an Easy-nLC 1200 System (Thermo Fisher Scientific) coupled to a Q Exactive HF mass spectrometer (Thermo Fisher Scientific) as described elsewhere (Kliza et al., 2017). The peptides were injected onto the column in HPLC solvent A (0.1% formic acid) at a flow rate of 500nl/min and subsequently eluted with a 87 min gradient of 10–33-50-90% HPLC solvent B (80% ACN in 0.1% formic acid). During peptide elution the

flow rate was kept constant at 200nl/min. The seven most intense precursor ions were sequentially fragmented in each scan cycle using higher energy collisional dissociation (HCD). The target values for the MS scan and MS/MS fragmentation were 3×10^6 and 10^5 charges with a maximum fill time of 25ms and 110ms, respectively. In all measurements, sequenced precursor masses were excluded from further selection for 30s.

The MS data were processed with MaxQuant software suite v.1.5.2.8 (Cox and Mann, 2008). Database search was performed using the Andromeda search engine (Cox et al., 2011), which is a module of the MaxQuant. MS/MS spectra were searched against a database consisting of 6,721 protein entries from *Saccharomyces cerevisiae*, the sequence of Cyc8p and a database consisting of 285 commonly observed contaminants. In database search, full tryptic specificity was required and up to two missed cleavages were allowed. Protein N-terminal acetylation, and oxidation of methionine were set as variable modifications. Initial precursor mass tolerance was set to 4.5ppm (parts per million), at the fragment ion level 20ppm was set for HCD fragmentation. Peptide, protein, and modification site identifications were filtered using a target-decoy approach at a false discovery rate (FDR) set to 0.01 (Elias and Gygi, 2007). For protein group quantitation a minimum of two quantified peptides were required.

SCP160 expression shut off by Tetracycline promoter

To deplete Scp160p, we used a strain in which *SCP160* expression was under the control of a tetracycline repressible promoter (Garí et al., 1997; Hirschmann et al., 2014). RJY4814 was inoculated into SC-Ura + 2% glucose + 2µg/mL doxycycline to an $OD_{600} = 0.08$ (approximately 8×10^5 cells/mL) and grown at 30°C with aeration for 6 hours. Subsequently, the cells were pelleted and resuspended in SC-Ura + 2% galactose + 1% raffinose + 2µg/mL doxycycline to an $OD_{600} = 0.15$ (approximately 1.5×10^6 cells/mL). Induction of 97QmCh occurred at 30°C with

aeration for 7 hours, at which point cells were fixed for observation of ploidy by flow cytometry or 97QmCh aggregation by microscopy as described in text.

Ploidy analysis by flow cytometry

Propidium iodide staining of DNA and flow cytometry was used to assess the ploidy of RJY4814 after repression of *SCP160* and induction of 97QmCh. Cells were fixed and treated as described in (Baum et al., 2004) and (Hirschmann et al., 2014). Propidium iodide fluorescence was detected on a Becton Dickinson Biosciences LSR II using FACSDiva version 6.1.3.

SUPPLEMENTAL REFERENCES

Baum, S., Bittins, M., Frey, S., and Seedorf, M. (2004). Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes. Biochem. J. *380*, 823–830.

Boersema, P.J., Geiger, T., Wisniewski, J.R., and Mann, M. (2013). Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. Mol. Cell. Proteomics *12*, 158–171.

Chen, D.-C., Yang, B.-C., and Kuo, T.-T. (1992). One-step transformation of yeast in stationary phase. Curr. Genet. *21*, 83–84.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J. V, and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods *4*, 207–214.

Garí, E., Piedrafita, L., Aldea, M., and Herrero, E. (1997). A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. Yeast *13*, 837–848.

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R., and Hofacker, I.L. (2008). The Vienna RNA websuite. Nucleic Acids Res. *36*, W70-4.

Halfmann, R., and Lindquist, S. (2008). Screening for amyloid aggregation by Semi-Denaturing Detergent-Agarose Gel Electrophoresis. J. Vis. Exp. 11–13.

Hirschmann, W.D., Westendorf, H., Mayer, A., Cannarozzi, G., Cramer, P., and Jansen, R.-P. (2014). Scp160p is required for translational efficiency of codon-optimized mRNAs in yeast. Nucleic Acids Res. *42*, 4043–4055.

Kliza, K., Taumer, C., Pinzuti, I., Franz-Wachtel, M., Kunzelmann, S., Stieglitz, B., Macek, B., and Husnjak, K. (2017). Internally tagged ubiquitin: a tool to identify linear polyubiquitin-modified proteins by mass spectrometry. Nat. Methods *14*, 504–512.

Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K., and Schiebel, E. (1999). Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. Yeast *15*, 963–972.

Longtine, M.S., McKenzie, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene

deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953–961.

Park, S.-H., Kukushkin, Y., Gupta, R., Chen, T., Konagai, A., Hipp, M.S., Hayer-Hartl, M., and Hartl, F.U. (2013). PolyQ proteins interfere with nuclear degradation of cytosolic proteins by sequestering the Sis1p chaperone. Cell *154*, 134–145.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.